

NPS ARCHIVE
1965.06
SKOLNICK, A.

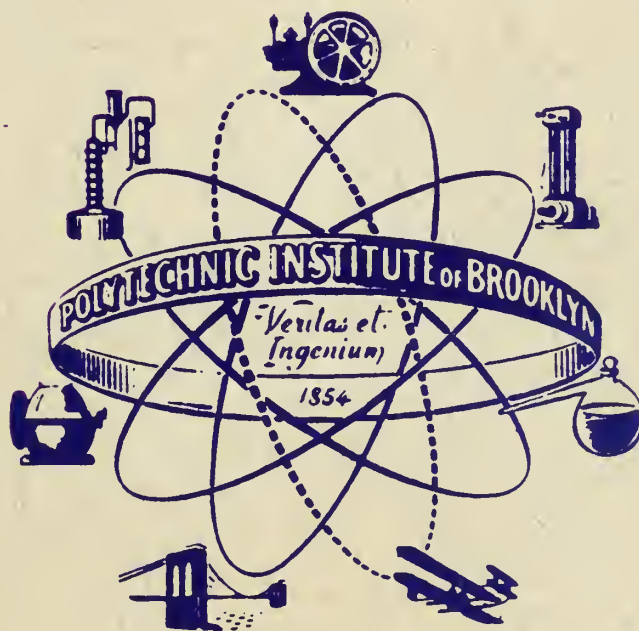
STABILITY AND PERFORMANCE OF MANNED CONTROL SYSTEMS

A Theory for Meeting the Stability and Performance Requirements of
Certain Man-Machine Systems

by

Alfred Skolnick

JUNE 1965



POLYTECHNIC INSTITUTE OF BROOKLYN
SYSTEM SCIENCE

Library
U. S. Naval Postgraduate School
Monterey, California

DUDLEY KNOX LIBRARY
NAVAL POSTGRADUATE SCHOOL
MONTEREY CA 93943-5101

STABILITY AND PERFORMANCE OF MANNED CONTROL SYSTEMS

A Theory for Meeting the Stability and Performance Requirements of
Certain Man-Machine Systems

DISSERTATION

Submitted in Partial Fulfillment
of the requirements for the
degree of

DOCTOR OF PHILOSOPHY (Systems)

at the

POLYTECHNIC INSTITUTE OF BROOKLYN

by

Alfred Skolnick

JUNE 1965

Approved:

19

Copy No. _____

William A. Lynch, Chairman
System Science Committee

Approved by the Guidance Committee:

Major: System Science

John G. Truxal
Professor of Electrical Engineering
Dean of Engineering

William A. Lynch
Professor of Electrical Engineering
Head of Department of Mechanical Engineering

Minor: Mathematics

Clifford W. Marshall
Associate Professor of Mathematics

Minor: Electrical Engineering

Joseph J. Bongiorno, Jr.
Associate Professor of Electrical Engineering

Microfilm or other copies of this dissertation are obtainable from the firm of

University Microfilms
313 N. First Street
Ann Arbor, Michigan

VITA

Alfred Skolnick was born on 15 August 1930 in Brooklyn, New York. He attended Brooklyn Technical High School and was graduated from Queens College in June 1951 with a Bachelor of Science degree in Mathematics. He received a Master of Arts degree in Mathematics from Columbia University in June 1952 and was employed as an Aerophysicist by Chance Vought Aircraft, Dallas, Texas until June 1953 when he enlisted in the United States Navy. Following graduation from Officer Candidate School in November 1953, he had duty at the Applied Physics Laboratory/Johns Hopkins University where he acted as Surface to Air Missile Project Officer. In 1955 he was accepted into the Regular Navy and became a career officer. In 1956 and 1957 he served in USS BOSTON (CAG-1), the world's first missile cruiser, as Fire Control Officer. In January 1960, he was granted the degree of Master of Science in Electrical Engineering (with distinction) from the U.S. Naval Postgraduate School, Monterey, California. He has served in various Navy Research and Development programs including the Polaris Program where he was Systems Integration Officer for Ships Inertial Navigation Systems. In 1962 he was selected for the Navy's new Science Ph. D. program. Concentrated thesis effort was begun in the Spring of 1964 and the study was completed in April 1965. The author is a member of Sigma Xi and presently holds the rank of Lieutenant Commander, United States Navy.

To both my wives:

The United States Navy - which paid for it
Sally - who made it possible

ACKNOWLEDGMENT

The difficulties associated with "going back to school" are many; those connected with a formal program leading to a doctorate in the engineering sciences are almost infinite. The problems are confounded when the "student" is a professional military officer whose normal responsibilities are somewhat removed from the abstracts and intangibles of Academe. To the United States Navy, I wish to express my gratitude for its willingness to experiment with a novel concept. Throughout the entire period required for study and severe concentrated effort, the Navy approach has been one of laissez-faire. This constructive attitude aided considerably in the successful attainment of the degree. To this end, Mrs. Clara Gumz of the Bureau of Naval Personnel, has been instrumental in developing the enlightened approach to naval officer education.

The opportunity for advanced study over a broad area also was contingent upon the flexibility and foresight of the Polytechnic Institute of Brooklyn and its Faculty in establishing a Systems program in advance of most educational institutions. The support provided by the Institute to certain of the administrative costs of the thesis is also sincerely appreciated.

Any ordinary acknowledgment to Professor John G. Truxal, Dean of Engineering, and my thesis adviser, is inadequate. His perspective on matters technical and otherwise has been remarkable; I am pleased to count him friend as well as teacher.

Professors Joseph J. Bongiorno, Jr., Ludwig Braun, Jr., Peter Dorato and Clifford W. Marshall of the Polytechnic Faculty provided considerable aid in their willingness to discuss, at all times, many of the concepts and technical points in this thesis. Their suggestions frequently rounded some of the procedural, and other, corners I sometimes backed myself into.

I am grateful to Mr. Lawrence Levey of the Polytechnic staff for his tol-

erance in acting as a sounding board and for his geometrical insight.

The Paul Moore Research and Development Center of the Republic Aviation Corporation, Farmingdale, Long Island, was kind enough to permit use of their analog computer, considerably facilitating the preparation and solution of the example in Chapter Six. For this courtesy, I particularly thank Mr. Ralph Hall, Manager, Dynamics Subdivision of the Space-Missiles Division and Mr. George Brandeau, Specialist Dynamics Engineer; Mr. Julius Port of the Analog Facility and Mr. Phillip Wall of Advanced Vehicle Systems were especially helpful in familiarizing me with the equipment.

Dr. Milton Katz, Head, Communications Psychology Division, U. S. Naval Training Devices Center, Port Washington, New York eased the burdens of initial library research by making available many reports in his files as well as in the collection of the Center Library.

The typing, as in many theses conducted under the grinding pressure of an absolute time deadline, was stalwartly performed by Miss Phyllis Incorvaia, who somehow was able throughout to maintain her sense of humor. Able typing assistance also was supplied by Mrs. Barbara Brown. The rigid time requirements could not have been met without the cooperation of Mrs. Helen Warren, Administrative Assistant at the Polytechnic's Long Island Graduate Center, who not only guided the clerical and reproduction effort, but "manned" the typewriter when the situation demanded. Mr. Sylvan Gebel, Graduate Center Librarian, diligently brought to my attention many recent studies connected with the human operator.

These expressions of appreciation would be incomplete without mention of my parents who always have held high the torch of education.

Finally, to my wife, Sally, and children, David and Susan (5 and 3), who have provided for me a "liberty port" during this three year endeavor while I have been "at sea", I solemnly promise never to do this again.

AN ABSTRACT

STABILITY AND PERFORMANCE OF MANNED CONTROL SYSTEMS

by

Alfred Skolnick

Adviser: John G. Truxal

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy (Systems)

The human-operator transform is usually specified as a function of the controlled-element characteristics and type of input power spectrum. Thus, any single description for the human operator really is valid only for a set of limited conditions. This study demonstrates the human-operator transfer function approach may be generalized; from a variety of published experimental data, "capability bounds" upon the transfer function parameters are formed. With such ranges defined, these parameters, forming a type of variable structure, can be used as a mapping function to display in the complex plane, the boundaries of human-adaptive capacity. These boundaries contain a collection of "critical points", similar to the $-1 + j0$ point on the usual Nyquist diagram; when the plant Nyquist contour is plotted in conjunction with these bounded areas (at particular frequencies), interpreta-

tion of system stability characteristics is possible. Several plots of such boundaries are presented for important frequencies in the human range. A man-monitored system configuration also appears to be susceptible to such stability analysis.

Experimental verification of the theory is obtained by analyzing certain systems which have been studied empirically in the past and for which results have been published in the literature. To this end, transparent overlays are provided which facilitate stability determination. The theoretical predictions parallel the experimental results remarkably well; one of the sets of data relates to a system which is open-loop unstable.

A discussion of scalar performance indices and some of their limitations is followed by a laconic survey of methods for solving the "best" performance problem, especially for varying-parameter plants. The notion of a "vector" performance index is introduced and a technique for establishing one response as "better" than another is presented. This technique rests intimately upon the formulation of an "artificial" function which, when reduced to zero, identically satisfies the upper bounds on system requirements. These requirements are set forth in the "vector" index. The design for a "best-overall" compensator to meet specifications upon a varying-parameter plant is obtained by using a "direct search" method (especially suited to the electronic computer). This method incorporates a "pattern strategy" to aid in finding the next set of trial values for the compensator; the strategy affords rapid convergence to a solution, as is demonstrated by an example.

Finally, the example itself displays the versatility of the variable-parameter mappings presented previously. Setting the human-operator transfer function equal to unity shows that the mappings have use for "inanimate" system design.

TABLE OF CONTENTS

| | Page |
|---|------|
| 1. STABILITY AND PERFORMANCE OF MANNED CONTROL | 1 |
| 1.1 Introduction and Historical Background | 1 |
| 1.2 Human-Operator Adaptive Capacity | 5 |
| 1.3 The "Difference Vector" Plot | 7 |
| 1.4 The Mapping Function and the Stability Criterion | 11 |
| 1.5 Applications of the Theory | 20 |
| 1.6 "Best" Performance and "Best-Overall" Compensation of Varying-Parameter Plants | 24 |
| 1.7 Summary | 34 |
| 2. THE HUMAN OPERATOR IN CONTROL SYSTEMS | 35 |
| 2.1 Introduction | 35 |
| 2.2 Controllability and Adaptation | 39 |
| 2.3 Thesis Organization | 41 |
| 2.4 The Quasi-Linear Model | 41 |
| 2.5 Adaptive Ranges of the Human Operator | 47 |
| 2.6 Summary | 52 |
| 3. HUMAN OPERATOR ADAPTIVE RANGES AND STABILITY THEORY | 53 |
| 3.1 Introduction | 53 |
| 3.2 System Transfer-Function Manipulation | 54 |
| 3.3 The Nominal Function | 56 |
| 3.4 Stability Criterion | 59 |
| 3.5 Summary | 64 |

| | Page |
|--|------|
| 4. THE STABILITY CRITERION AND THE VARIABLE PARAMETER MAPPING | 65 |
| 4.1 Introduction | 65 |
| 4.2 Analysis of a Simple System | 65 |
| 4.3 Multidimensional-Variable-Parameter Mapping | 84 |
| 4.4 A Map of the Human-Operator Adaptive Range | 88 |
| 4.5 Summary | 108 |
| 5. APPLICATIONS OF THE THEORY | 109 |
| 5.1 Introduction | 109 |
| 5.2 The Hall and Smith Studies | 109 |
| 5.3 The Hall Study | 110 |
| 5.4 The Smith Study | 127 |
| 5.5 Summary | 135 |
| 6. PERFORMANCE | |
| 6.1 Introduction | 137 |
| 6.2 Varying-Parameter Plants and "Best" Performance | 139 |
| 6.3 The Concept of a "Vector" Performance Index | 147 |
| 6.4 The Gradient Method | 153 |
| 6.5 Direct Search and Pattern Strategy | 154 |
| 6.6 "Best-Overall" Compensation of a Varying-Parameter System - An Example | 158 |
| 6.7 Summary | 184 |
| APPENDIX A | 185 |
| A.1 Human-Operator Zero Locus | 185 |

| | Page |
|---|------|
| APPENDIX B | 187 |
| B.1 The Parabolic Effect | 187 |
| B.2 An Exploration of T_L, T_I, T_N Space | 195 |
| APPENDIX C | 204 |
| C.1 A Two Parameter System | 204 |
| C.2 The Degenerate Mapping | 206 |
| BIBLIOGRAPHY | 212 |

LIST OF FIGURES

| Figure | Page |
|--|------|
| 1-1 The Human Controller in the Compensatory Mode | 3 |
| 1-2 The Human Controller as a Monitor | 8 |
| 1-3a Plot of $-L_n$ and Ω | 12 |
| 1-3b The "Difference Vector" Plot | 12 |
| 1-4 Human-Operator Adaptive Range | 13 |
| 1-5 Typical Human-Operator Variable-Parameter Mapping ($\omega = 1$) | 15 |
| 1-6 Variable-Parameter Mapping ($\omega = 0.5$) | 16 |
| 1-7 Variable-Parameter Mapping ($\omega = 1.0$) | 17 |
| 1-8 Variable-Parameter Mapping ($\omega = 2.0$) | 18 |
| 1-9 Variable-Parameter Mapping ($\omega = 3.0$) | 19 |
| 1-10 Plot of Configuration C21 | 22 |
| 1-11 Human Control with "Best" Compensation | 25 |
| 1-12 Compensation of a Varying-Parameter System | 29 |
| 1-13 "Direct Search" with "Pattern Strategy" | 32 |
| 2-1 The Human Controller in the Compensatory Mode | 43 |
| 2-2 The Human Controller as a Monitor | 51 |
| 3-1 Separation of Fixed and Varying-Parameter Elements | 57 |
| 3-2a Plot of $-L_n$ and Ω | 62 |
| 3-2b The "Difference Vector" Plot | 62 |
| 4-1 A System with One Variable-Parameter | 66 |
| 4-2 Stable Closed-Loop | 67 |
| 4-3 Unstable Closed-Loop | 67 |
| 4-4 Effect of Variable-Parameter | 67 |
| 4-5 Stability Range Display | 67 |

| Figure | | Page |
|--------|---|------|
| 4-6 | Dual Nyquist Plot of a Varying-Parameter System | 73 |
| 4-7 | Detailed Dual Nyquist Plot | 76 |
| 4-8 | Hypothetical Crossing Situations (in the small) | 77 |
| 4-9 | Hypothetical Crossing Situations (in the large) | 79 |
| 4-10 | Excursion of the Difference Vector | 80 |
| 4-11 | Hypothetical Crossing Situation (in the large) | 81 |
| 4-12 | Excursion of the Difference Vector | 82 |
| 4-13 | Boundary Mapping Resulting in Complete Hull | 86 |
| 4-14 | Boundary Mapping with Hull which is Not Complete (also shows Convex Hull) | 86 |
| 4-15 | Hypothetical Variable-Parameter Situation | 87 |
| 4-16 | Human-Operator Adaptive Range | 89 |
| 4-17 | Mapping of Human-Operator Adaptive Range | 90 |
| 4-18 | Typical Human-Operator Variable-Parameter Mapping ($\omega = 1$) | 96 |
| 4-19 | Variable-Parameter Mapping ($\omega = 0.5$) | 102 |
| 4-20 | Variable-Parameter Mapping ($\omega = 1.0$) | 103 |
| 4-21 | Variable-Parameter Mapping ($\omega = 2.0$) | 104 |
| 4-22 | Variable-Parameter Mapping ($\omega = 3.0$) | 105 |
| 4-23 | Variable-Parameter Mapping ($\omega = 10.0$) | 106 |
| 4-24 | The Convex Hull | 107 |
| 5-1 | Plot of Configuration C21 | 115 |
| 5-2 | Plot of Configuration C32 | 116 |
| 5-3 | Plot of Configuration C30 | 123 |
| 5-4 | Plot of Configuration C25 | 124 |
| 5-5 | Plot of Configuration C31 | 125 |
| 5-6 | Open-Loop Unstable Plant ($\zeta = -0.7$, $\omega_n = 2.5$) | 128 |

| Figure | | Page |
|--------|---|------|
| 5-7 | Open-Loop Unstable Plant ($\zeta = -0.35, \omega_n = 7.2$) | 129 |
| 5-8 | Interpreting Open-Loop Instability Effects | 132 |
| 6-1 | Human Control of Varying-Parameter Plant | 142 |
| 6-2 | Human Control of Varying-Parameter Plant | 142 |
| 6-3 | Human Control of Varying-Parameter Plant | 142 |
| 6-4 | Solutions Outside the Region of Stability | 146 |
| 6-5 | Human Control with "Best" Compensation | 148 |
| 6-6 | Stability and Performance Regions in Transform Domain (a) and Time Domain (b) | 152 |
| 6-7 | Perturbation Pattern for Gradient Determination | 155 |
| 6-8 | Compensation of a Varying-Parameter System | 157 |
| 6-9 | Sketch of $-L_n$ Plots | 157 |
| 6-10 | Controlled Element Plots, $-L_n$, for various ζ | 160 |
| 6-11 | Analog Computer Setup for Direct Search Example | 164 |
| 6-12 | "Direct Search" with "Pattern Strategy" | 165 |
| B-1 | Typical Map from (T_L, T_I, T_N) Space | 188 |
| B-2 | Scanning a Face in (T_L, T_I, T_N) Space | 189 |
| B-3 | The Parabolic Effect | 190 |
| B-4 | Effect of Varying T_L | 196 |
| B-5 | Effect of Reduced Range on T_I | 202 |
| C-1 | Second-Order System with Compensator | 205 |
| C-2 | The Degenerate Three-Parameter Mapping for Two-Parameter Analysis | 207 |
| C-3 | The Degenerate Map and $-L_n = \frac{-\omega_n^2}{s(s + 2\zeta\omega_n)}$ | 208 |
| C-4 | Two-Parameter Variation for $\alpha = 2, \beta = 25$; also shows $-L_n$ | 210 |

"Instead, we shall regard all that has been presented so far in this volume, and all that will follow, as merely a set of conceptual and mathematical tricks, a bit of scientific legerdemain, which, for reasons unknown, but for which we are profoundly grateful, prove to be startlingly more successful in many applications than anyone who examines the basic assumptions would have any right to expect."

Richard Bellman
Adaptive Control Processes:
A Guided Tour

CHAPTER ONE

STABILITY AND PERFORMANCE OF MANNED CONTROL SYSTEMS1.1 Introduction and Historical Background

The unique capabilities of man, while seeming to pale in the face of increasingly sophisticated and automatic equipment, still offer characteristics either wholly unattainable by machines or, at the least, constrained in their attainment by weight, reliability, or economic considerations. Ultimately, the purpose of any good system, e. g., economic, political or control, is to ease the burdens of man, either directly (automatically serving his convenience, transporting him to some preferred location), or indirectly (aiding him in attaining his desires, whether the production of goods or the retention of vehicle control). It is not surprising, therefore, to find that a good deal of study and experiment have been devoted to various aspects of human behavior. These range from investigations of a psychological variety to those of a less ambitious nature involving the characterization of human dynamics. Even in the latter, less extensive domain, any attempt at describing the rich diversity of human dexterity is an ambitious one. To permit mathematical description of human response to stimuli, the scope of the problem must be reduced (to a tracking task with visual inputs) and the input-signal characteristics constrained to certain classes. Using such techniques, Tustin, working on the problem of man-operated-gun-fire-control equipment, first attempted to describe the human controller in servomechanism terms.⁽¹⁾ About that time, Sobczyk⁽²⁾ and

⁽¹⁾ Tustin, A., An Investigation of the Operator's Response in Manual Control of a Power-Driven Gun, C.S. Memo 169, Metropolitan-Vickers Co., Ltd., Sheffield, England, 1944.

⁽²⁾ Sobczyk, A., Aided Tracking, MIT Rad. Lab. Rpt. 430, Sept., 1943.

Phillips⁽¹⁾ performed similar work on "aided" gun-laying systems*. It was Tustin, however, who first arbitrarily separated the input-output characteristics of man into a linear operator term and a "remnant" term, to account for that part of the human-output response not attributable to the linear operation on the input, Fig. 1-1. It is clear also from Fig. 1-1 that the system under consideration restricts the human controller to processing the error signal; this configuration is called "compensatory tracking", a term assigned by the engineering psychologists. Certainly, not all human control can be classified as tracking, nor does the tracking task imply human control. A sufficiently large group of situations is covered by the compensatory tracking mode, however, to make its study very worthwhile; even in those situations where, due to other complications, this mode does not provide an accurate description of system conditions, an understanding of such manual-tracking capabilities can provide insight into the more complex situation.

It is evident that a mechanism so adaptable that it can "shoot marbles" or pilot high performance aircraft does not lend itself to description by a single transfer function. After World War II, several now-classic studies of the human controller were performed.⁽²⁾⁽³⁾ From these and the comprehensive work of McRuer and Krendel, there began to evolve a general form of human-operator-transfer-characteristic in which some confidence could be placed.⁽⁴⁾ It became apparent that, in general, the particular transfer description was a function of both the controlled element and the input power spectrum. Thus,

*In aided tracking, the human controller's output signal directly affects both the position and rate of the gun; thus, it is possible to simultaneously correct errors in gun-mount displacement and speed.

(1) Phillips, R.S., H.M. James, and N.B. Nichols, "Theory of Servomechanisms", McGraw-Hill, New York, 1947.

(2) Russell, L., Characteristics of the Human as a Linear Servo Element, M.S. Thesis, MIT, May, 1951.

(3) Elkind, J.I., Characteristics of Simple Manual Control Systems, Ph.D. Thesis, MIT, April, 1956; also issued as MIT Lincoln Lab. Rep. No. 111.

(4) McRuer, D.T. and E.S. Krendel, Dynamic Response of Human Operators, WADC Technical Rept. 56-524, Oct., 1957.

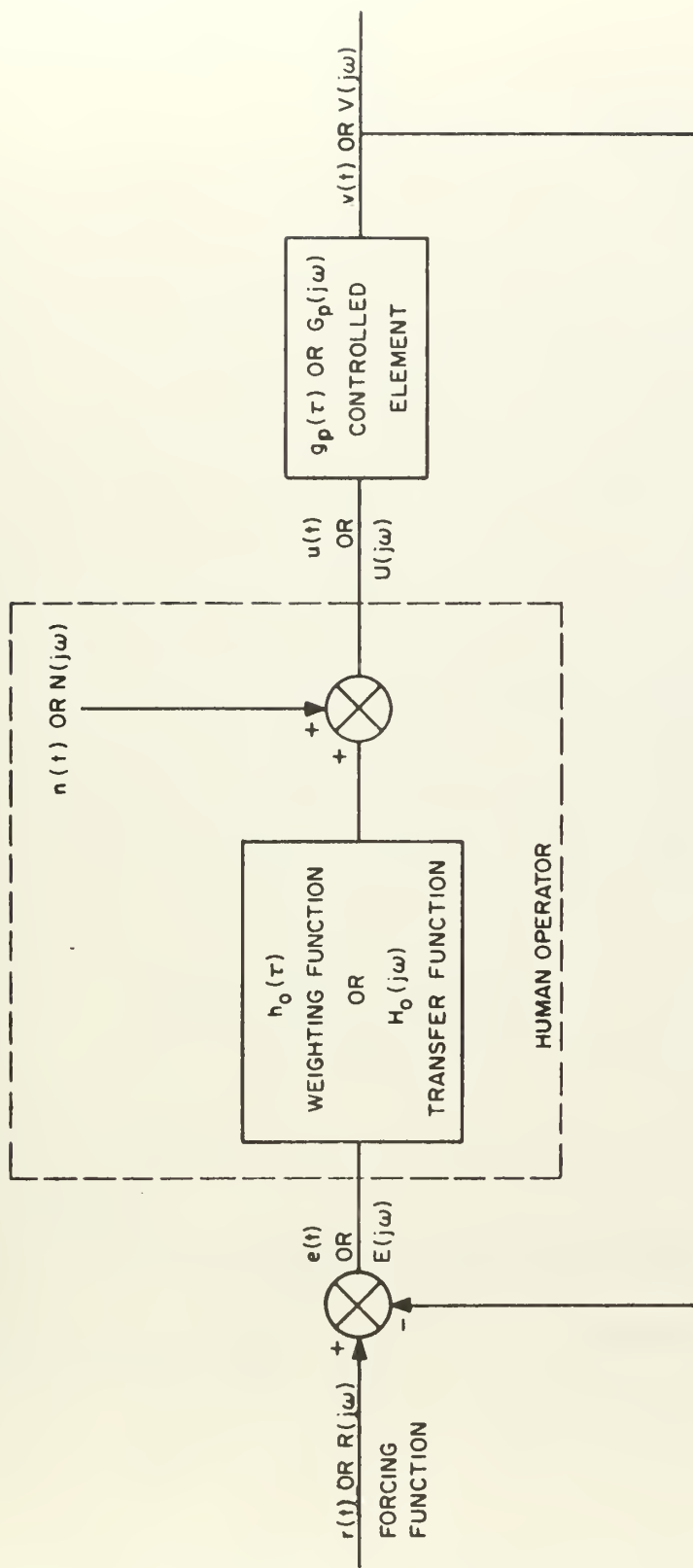


Fig.1-1 : The Human Controller in the Compensatory Mode

combinations of useful sets of plant dynamics and input power spectra were presented in a table with certain "adjustment rules" which allowed the designer to "set" the parameters in the human-operator transfer function in accordance with empirical evidence. As a result of the environmental dependence of the human-operator transfer function and the occasional lack of point-to-point correspondence between the model and actual tracking records, the term "quasi-linear" is used to qualify the transfer function description; thus the model portrays a smoothed, time-averaged type of "linearity". Since the publication of the McRuer compendium, additional studies have demonstrated further the utility of the transfer function approach.⁽¹⁾⁽²⁾ The assumption of human operator stationarity^{*} is inherent in most of the studies involving the transfer function approach, although non-steady behavior is, to some extent, measured by the remnant term; work has been conducted to determine time-varying descriptions of the human controller.⁽³⁾⁽⁴⁾ Although many elegant models have been presented in the last few years, none has been tested and established to the extent of the model contained in the studies of McRuer and Krendel. This quasi-linear mathematical description, in its most general form, appears as

$$H_o(j\omega) = \frac{K_h e^{-j\omega T} (1+j\omega T_L)}{(1+j\omega T_N)(1+j\omega T_I)} \quad (1-1)$$

* That is, during any given measurement run, one assumes that human response is not time dependent.

- (1) Hall, I.A.M., Effects of Controlled Element on the Human Pilot, Princeton Univ., WADC Tech. Rept. No. 57-509, Aug., 1958.
- (2) Smith, R.H., On the Limits of Manual Control, IEEE Trans. on Human Factors in Electronics, Sept., 1963.
- (3) Sheridan, T.B., Studies of Adaptive Characteristics of the Human Controller Tech. Rept. No. ESD-TDR 62-351, L.G. Hanscom Field, Bedford, Mass., Dec., 1962.
- (4) Wierville, W.W., G.A. Gagne, A Theory for the Optimal Deterministic Characterization of the Time-Varying Dynamics of the Human Operator, Cornell Aero. Lab., NASA Contractor Rept. CR-170, Feb., 1965.

It is noted that the description contains a gain term, a pure lag (frequently associated with human reaction time), a pole with time constant T_N (taken to represent neuromuscular lag), and an equalization characteristic with time constants T_L and T_I . This equalization capability represents the human operator's capacity to "tune" his response in much the same manner a control engineer compensates a system by introducing lead or lag networks. Considerable experiment has shown that descriptions more complicated than (1-1) offer little added accuracy in "goodness of fit" to the data; frequently, however, the form of (1-1) can be simplified by eliminating some of the time constants whenever they are not required to obtain a good-fit to actual human-operator response data. In part, this represents the analytical analog of the human-operator adaptive capacity. Equation (1-1) is obtained by analysis of power spectra records; the expression is derived under the assumptions of time-invariance and "trained" operators, i.e., the human-controller transfer characteristic represents post-learning period response-capability. This assumed familiarity with the task introduces sufficient response-consistency to permit mathematical description of the data.

1.2 Human-Operator Adaptive Capacity

The transfer function approach continues to offer the most promise to engineers involved with human-controlled systems. Studies designed to increase the utility of this technique are still being pursued.⁽¹⁾⁻⁽⁵⁾ This thesis presents

-
- (1) Muckler, F.A. and R.W. Obermayer, The Use of Man in Booster Guidance and Control, NASA CR-81, July, 1964.
 - (2) Muckler, F.A. and R.W. Obermayer, Control System Lags and Man-Machine System Performance, NASA CR-83, July, 1964.
 - (3) Summers, L.G., and K. Ziedman, A Study of Manual Control Methodology with Annotated Bibliography, NASA CR-125, Nov., 1964.
 - (4) Bergeron, H.P., J.K. Kincaid, and J.J. Adams, Measured Human Transfer Functions in Simulated Single-Degree-of-Freedom Nonlinear Control Systems, NASA TN D-2569, Jan., 1965.
 - (5) Obermayer, R.W. and F.A. Muckler, On the Inverse Optimal Control Problem in Manual Control Systems, NASA CR-208, April, 1965

a different approach to meeting the stability and performance requirements of certain man-machine systems. In the many studies conducted, there appear many different values for the parameters K_h , τ , T_L , T_I and T_N of (1-1); these distinct values indicate the "quasi-linear" character of the human-operator transfer function; also they simultaneously imply the adaptive capacity of the human controller. From these empirically derived values, "capability bounds" upon the human operator's adaptive range can be formed. Certain mathematical manipulations and a suitable method for displaying such bounds, when used in conjunction with plots of the controlled-element dynamics, permit stability estimates to be made for the combined closed-loop system. The assumption that the human operator regards system stability as a primary criterion is intrinsically contained in these estimates; it has been established in experiments that this is a valid supposition. Experiments also have shown that the human controller is most satisfied with the dynamics of a system when he is able to participate (and maintain control) as a simple amplifier.⁽¹⁾⁽²⁾ This property implies that any "margin of human capacity" remaining can serve as a measure of human satisfaction with the system ("handling qualities"); these margins appear as a part of the "capability bounds" display. The bounds formulated and used in this thesis cannot be regarded as definitive parameter ranges for large populations since such do not now exist. Indeed, if past experiments share any common feature, it is the limited number of subjects used to provide data. Usually, the small range of variation among well-trained subjects is cited as justification for limited subject sample size. The cross-sectional aspect of this thesis, i. e., its use of the data from many experiments, to some extent

(1) Birmingham, H. P. and F. V. Taylor, A Human Engineering Approach to the Design of Man-Operated Continuous Control Systems, NRL Rept. 43333, April, 1954.

(2) McRuer, D. T. and E. S. Krendel, The Human Operator as a Servo System Element, J. Franklin Inst., Vol. 267, Nos. 5 and 6, May and June, 1959.

attenuates the effect of reduced sample size. The ranges established for the parameters of (1-1) are as follows:

$$K_h : 0.6 - 250 \quad (1-2a)$$

$$\tau : 0.1 - 0.4 \quad (1-2b)$$

$$T_L : 0 - 5.3 \quad (1-2c)$$

$$T_I : 0 - 25 \quad (1-2d)$$

$$T_N : 0 - 0.7 \quad (1-2e)$$

The units for K_h depend upon the units of the input-signal to the human controller, the others have units of seconds. It should be noted that the ranges on K_h and τ do not restrict the analysis in any way; as is seen below, these two parameters enter into the method in such a way that any range desired can be accommodated quickly.

1.3 The "Difference Vector" Plot

For the system of Fig. 1-1 the closed-loop system function is^{*}

$$\frac{V}{R} = \frac{H_o G_p}{1 + H_o G_p} \quad (1-3)$$

The system is stable if the denominator contains no zeros in the right-half plane. It is noteworthy that in a monitoring configuration,^{**} Fig. 1-2, the closed-loop function is

$$\frac{V}{R} = \frac{H_o G_p + G_c G_p}{1 + H_o G_p + G_c G_p} \quad (1-4)$$

* Equation (1-3) does not consider the effect of $n(t)$. The stability analysis which follows does not depend upon this exclusion.

** In a monitoring configuration, the human operator observes system error and enters the loop as a "signal processor" only after he decides his contributions can improve system response. To include the "no entry" condition the lower bound on K_h (1-2a) is reduced to zero.

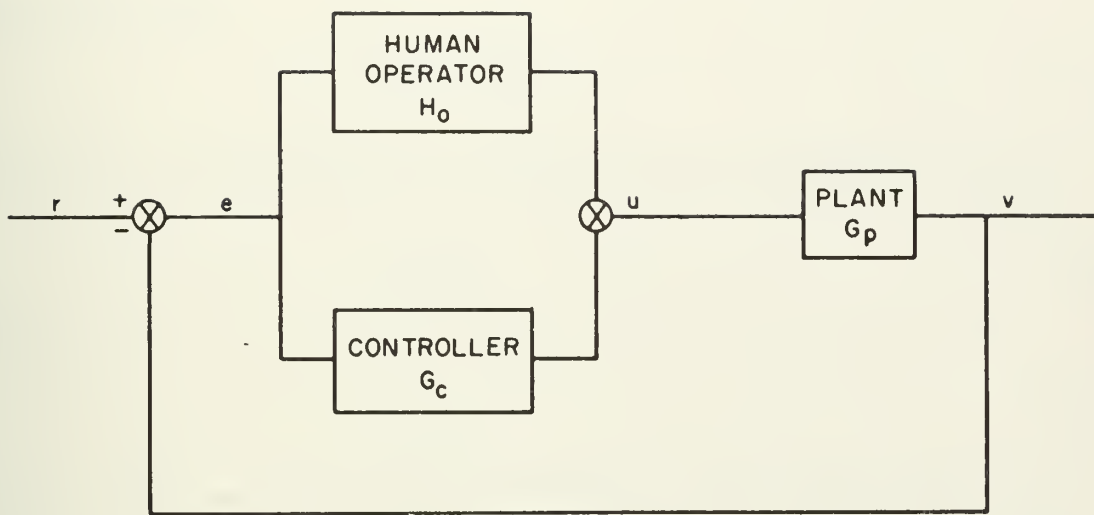


Fig. 1-2 : The Human Controller as a Monitor

Again, the system is stable only if the denominator contains no right-half plane zeros. The denominator may be written

$$(1 + G_c G_p) \left(1 + \frac{H_o G_p}{1 + G_c G_p} \right) \quad (1-5)$$

If

$$G'_p = \frac{G_p}{1 + G_c G_p} \quad (1-6)$$

then (1-5) becomes

$$(1 + G_c G_p)(1 + H_o G'_p) \quad (1-7)$$

In (1-7), the term containing the variable-parameter portion of the system is

$$1 + H_o G'_p \quad (1-8)$$

The term

$$1 + G_c G_p \quad (1-9)$$

in (1-7) presents no serious problem in stability analysis since it has a fixed-parameter structure. It should be noted that (1-8) has the form of the denominator of (1-3); of course, any effects from the poles of (1-6) remain to be considered. Expression (1-8), set equal to zero, is the characteristic equation for the system; for stability this equation must have no roots in the right-half plane.

$$1 + H_o G'_p = 0 \quad (1-10)$$

Substituting (1-1) into (1-10) and manipulating, one obtains

$$\frac{(1 + j\omega T_N)(1 + j\omega T_I)}{(1 + j\omega T_L)} + K_h e^{-j\omega\tau} G'_p = 0 \quad (1-11)$$

In Eq. (1-11), if the first term is called, Ω , and the second, L_n , then (1-11) becomes

$$\Omega + L_n = \frac{(1+j\omega T_N)(1+j\omega T_I)}{(1+j\omega T_L)} + K_h e^{-j\omega\tau} G'_p = 0 \quad (1-12)$$

By an extension of the Nyquist criterion,⁽¹⁾ the encirclements that $-L_n$ makes of the Ω region provide stability information about the system.

Usually, in a fixed-parameter problem the system structure depends for its stability upon the nature of encirclements of the "minus one" point by its Nyquist contour; with a varying parameter structure, this "critical point" generalizes into a "region of critical points", each of which is obtained from a specific set of values assigned to the triple (T_L, T_I, T_N) . Unfortunately in order to display this "region of variation" for the ranges on T_L, T_I, T_N , the term Ω of (1-12) must be plotted for particular frequencies $\omega = \omega_i$. Thus, to interpret "encirclements" and infer stability characteristics of (1-12), several sets of such "regions of variation" must be plotted for pertinent frequencies. Fortunately, in the human controller problem only a very few frequencies are required as a consequence of the rapid deterioration, at frequencies above $\omega \approx 20$ radians per second, of both the human operator and the accuracy of the mathematical model. It should be noted that (1-12) is written such that Ω represents the VARYING portion of the system and L_n the FIXED portion. In plotting these to determine stability (discussed below), the effects of varying-human-operator gain is readily observed on the Nyquist curve for $-L_n$; the human reaction time delay is handled by selecting pertinent values for τ and viewing their overall effects. Any number of such τ values can be included quickly and easily, since an added phase

⁽¹⁾Satche, M., Discussion, J. Appl. Mech., p. 419, Dec., 1949.

shift to $-L_n$ is all that is required.

Rather than plot the sum (1-12) and observe any rotations about the origin, the following technique is used. The negative of L_n is plotted, i.e., $-L_n$, as well as the region of variation Ω ; Fig. 1-3(a) displays a hypothetical situation. If a vector is drawn from the $-L_n(j\omega)$ locus to the Ω "locus", each end of the vector being at the same value of ω , that "difference vector" is actually the function $\Omega + L_n$ and the rotations it experiences give stability information for that function. Thus, stability is determined from the nature of the origin encirclements, if any, that occur in the "difference vector" plot, Fig. 1-3(b). The vectors shown in Fig. 1-3(a) are for specific values of the varying parameters as well as for particular frequencies. Thus, the locus of (1-12) may be visualized without carrying out the summation. The replacement of Ω by unity results in an ordinary Nyquist plot with the head end of the vector fixed at $+1 + j0$ and the other tracing out the $-L_n(j\omega)$ locus. Thus, as a dividend the plot of $-L_n$ allows immediate observation of the plant-stability characteristics in those cases when $G_c = 0$; the critical point from which stability is determined in such instances is at $+1 + j0$.

1.4 The Mapping Function and the Stability Criterion

With the human operator's gain-varying capability and delay time lumped into the fixed-parameter structure, the varying portion of the characteristic equation is given by

$$\Omega = \frac{(1+j\omega T_N)(1+j\omega T_I)}{(1+j\omega T_L)} \quad (1-13)$$

It is shown in Chapter Four that for fixed ω , T_N , and T_I , variations in T_L map into circular arc segments, while with the other parameters fixed, variations in T_I or T_N map into straight lines. In this way, the bounded region of variation for the ranges selected (1-2, c through e), Fig. 1-4, may be plotted. A typical

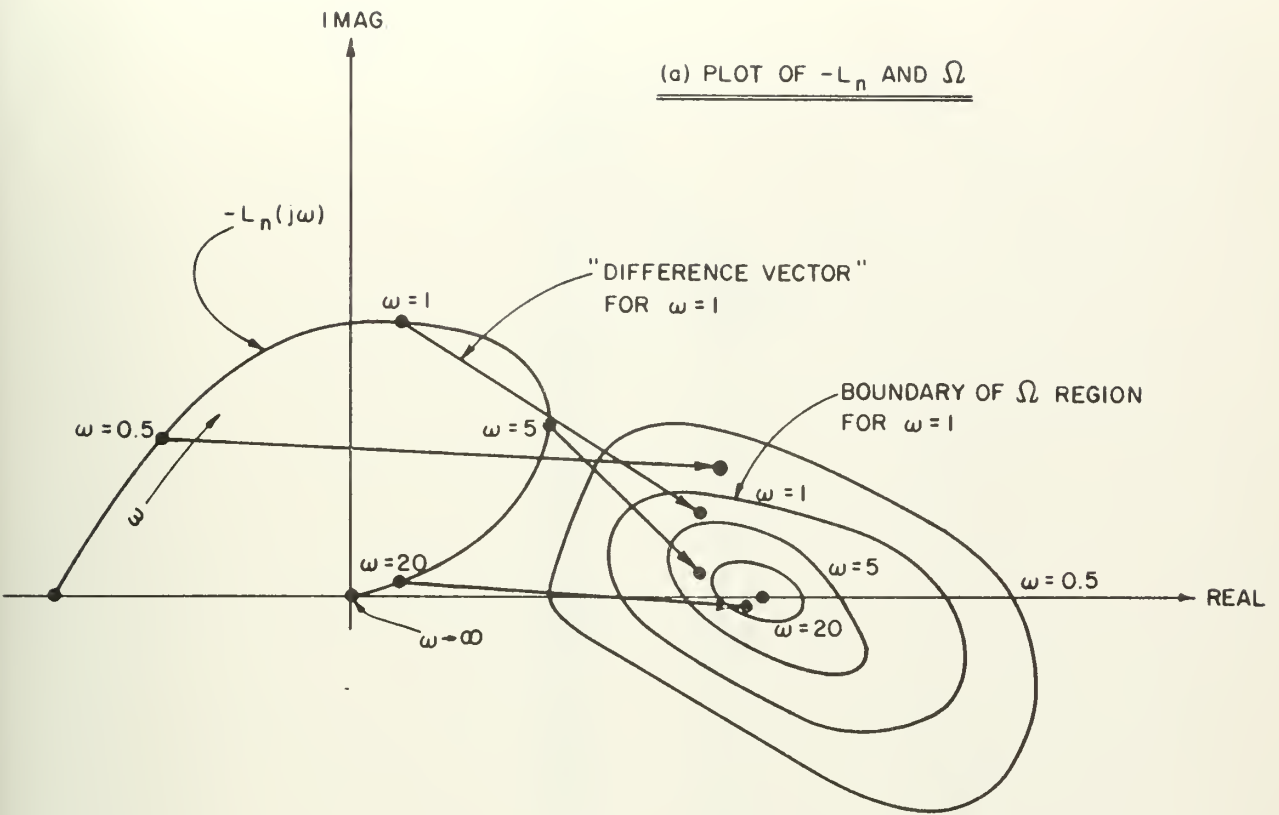
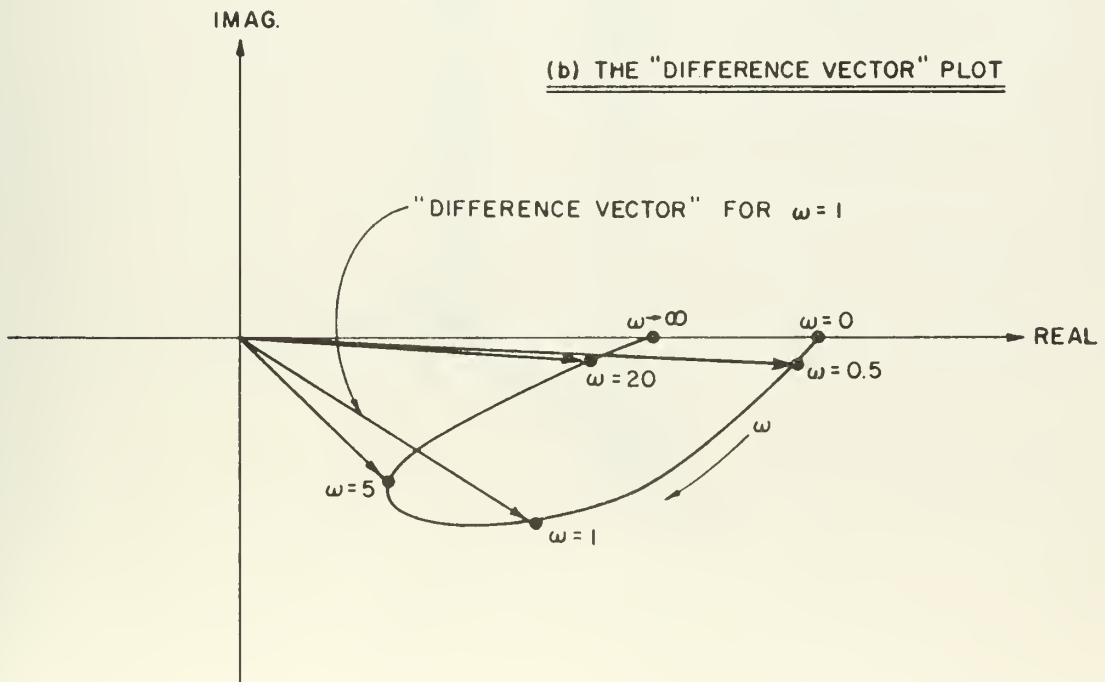
(a) PLOT OF $-L_n$ AND Ω (b) THE "DIFFERENCE VECTOR" PLOT

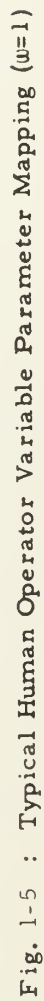
Fig. 1-3

plot appears in Fig. 1-5; this plot is a mapping through Ω of the boundaries (end points) of Fig. 1-4. Frequently, and in the case of (1-13), the "end point" mapping is not "complete". The property of completeness signifies that every point of the mapping Ω , as the space of points in Fig. 1-4 is scanned, is also a point of the "end point" mapping. This property is required if rigorous statements regarding stability criteria are to be made; i.e., completeness permits the establishment of necessary and sufficient conditions for stability. For human-controller studies, it may be shown that the mapping of Fig. 1-5 is "almost complete" and its application to engineering design studies is acceptable. Plots of Ω for the frequencies $\omega = 0.5, 1.0, 2.0$ and 3.0 are given in Figs. 1-6 through 1-9; these plots are "universal" in the sense that they may be used with any system of interest.

With the critical points governing stability defined by a region Ω , estimates of human-operator handling capability are made in the following way: if there is some area of Ω unenclosed by the contour, $-L_n$, then presumably the human operator can assume such parameter values and stabilize the system. The parameter values necessary to accomplish this appear to give insight into the amount of effort the human operator must expend to retain control.⁽¹⁾ If the entire Ω region is enclosed, the human operator cannot control under any circumstances (for the capability ranges assumed and plotted). If no portion of the Ω region is enclosed, there is stability no matter what the actions of the human operator (within the sense of the transfer function description).

The ultimate justification for this approach is provided by predicting whether the human operator can control (stabilize) certain systems. The systems used in this thesis were analytically equivalent to those reported on in

(1) McRuer, D.R., I.L. Ashkenas and C.L. Guerre, A Systems Analysis View of Longitudinal Flying Qualities, WADD 60-43, Wright Air Development Division, Jan., 1960.



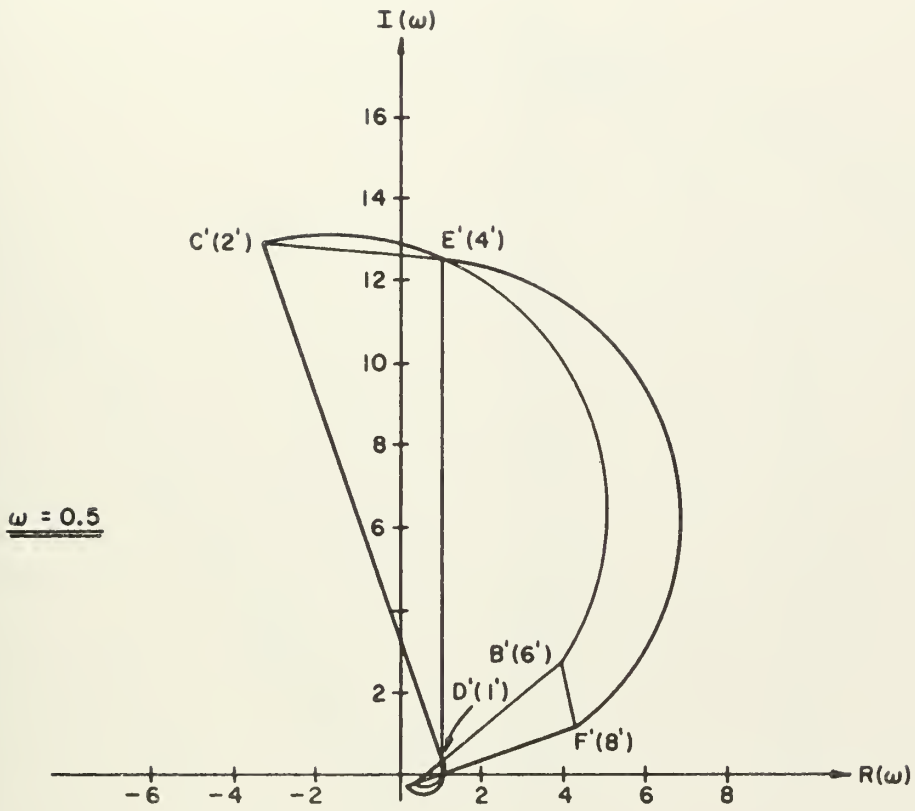


Fig. 1-6 : Variable Parameter Mapping

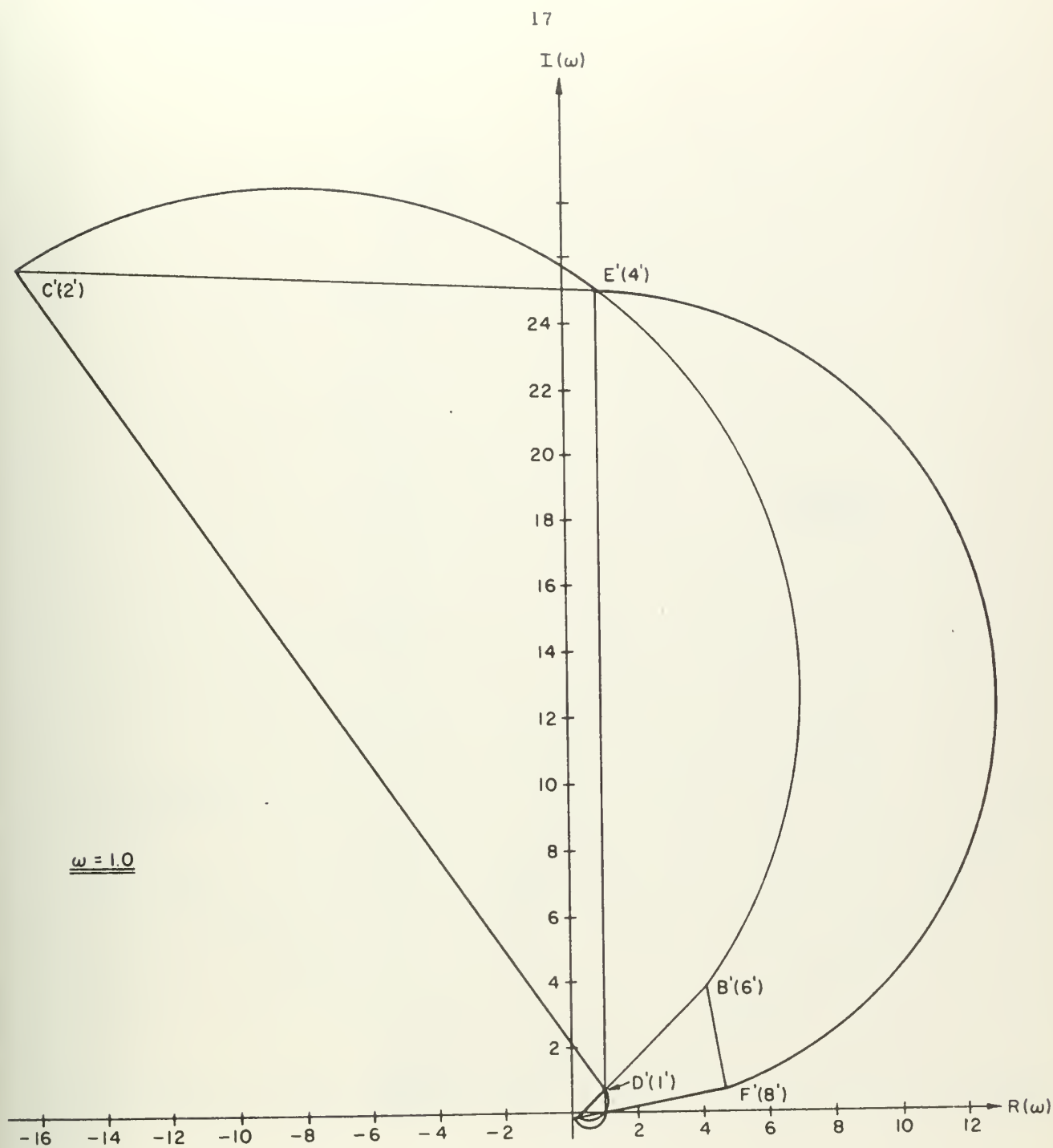


Fig.1-7 : Variable Parameter Mapping

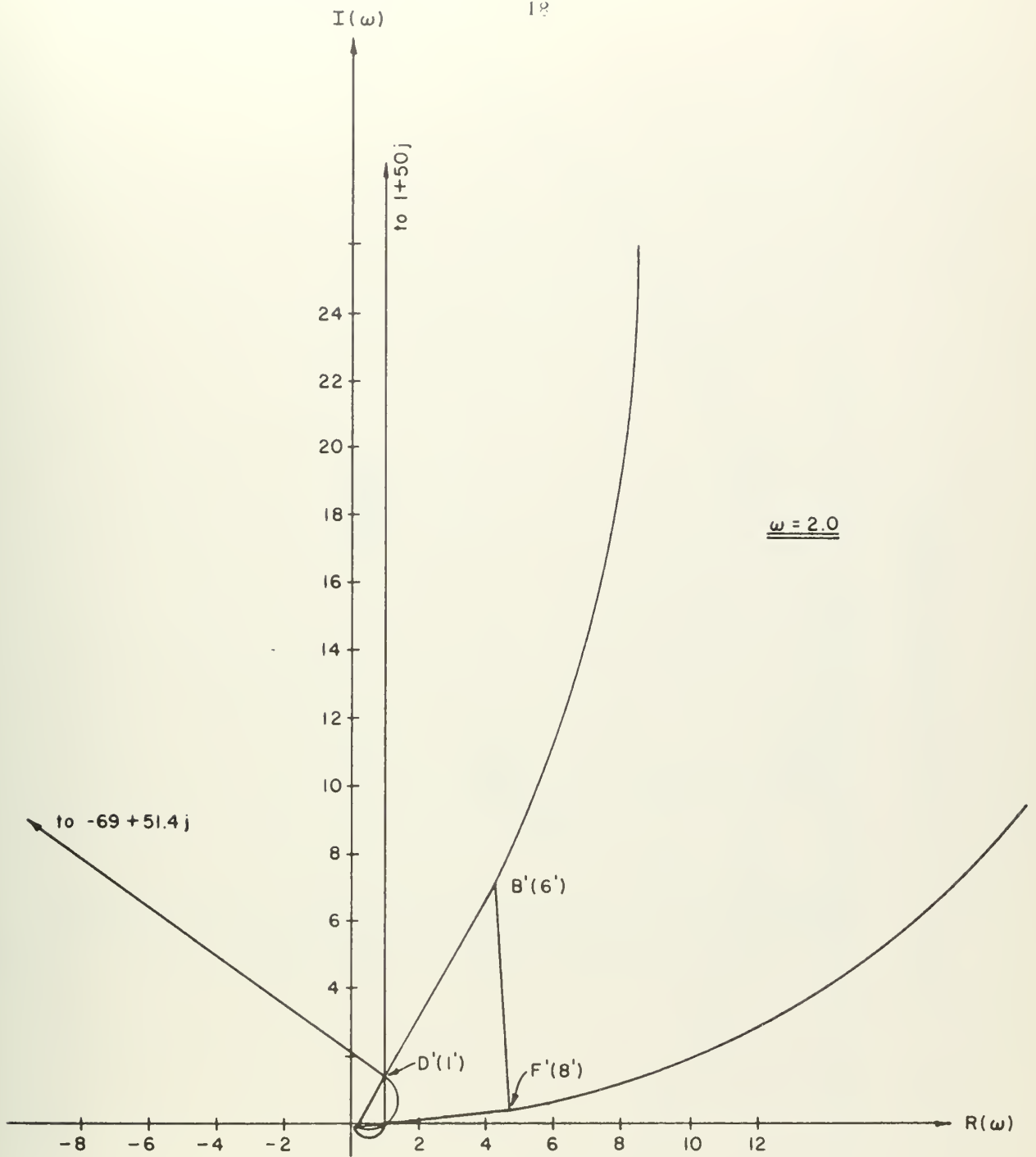


Fig. 1-8 : Variable Parameter Mapping

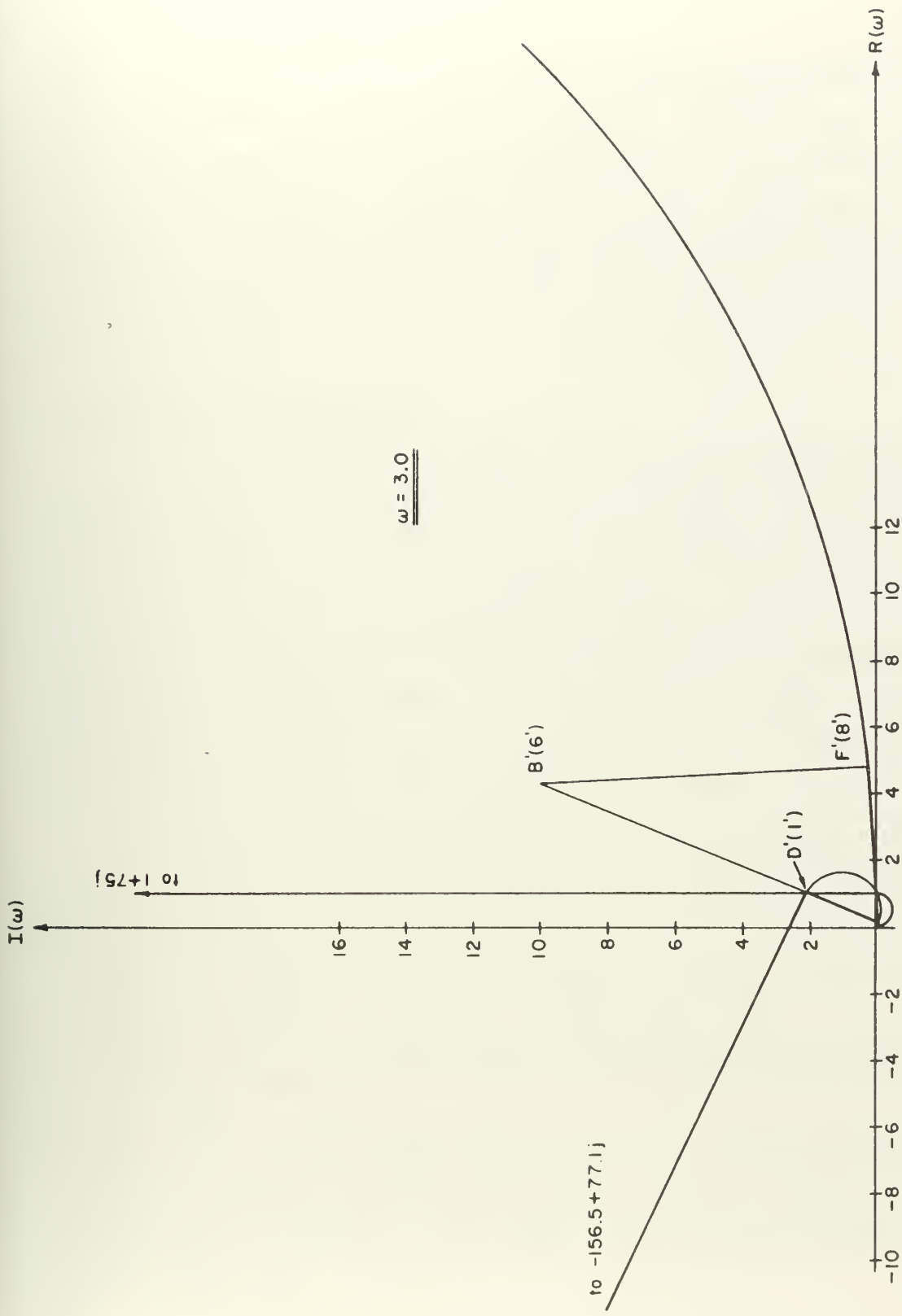


Fig. 1-9 : Variable Parameter Mapping

the literature. The published results were determined experimentally; their classifications were obtained from human operator evaluations (pilot opinion of handling qualities). The a priori predictions parallel the empirical results very well, including some for systems which are open-loop unstable. Qualitative information concerning handling qualities and engineering insight into how the systems may be modified to improve their stability and the human operator's satisfaction with their characteristics is also provided.

1.5 Applications of the Theory

In a simulator study by Ian Hall at Princeton University the effects upon a pilot of various controlled-element longitudinal dynamics were investigated.⁽¹⁾ McRuer et al performed a detailed analysis on the results of the Hall report in an attempt at developing a "flying qualities" theory.⁽²⁾ The McRuer approach follows the "adjustment rules" procedure described in Section 1.1. Another experimental study conducted at the University of Cincinnati by Ralph Smith investigated the maximum capabilities of a human in controlling certain open-loop-unstable systems.⁽³⁾ The results of both these studies furnish experimental justification for the theoretical predictions in this thesis.

As an indication of the analysis technique, a single example from the Hall study is offered in this chapter, other applications appear in Chapter Five. The experiment was conducted in the compensatory arrangement as shown in Fig. 1-1

-
- (1) Hall, I.A.M., Effects of Controlled Element on the Human Pilot, Princeton University, WADC Tech. Rept. No. 57-509, Aug., 1958.
 - (2) McRuer, D.T., I.L. Ashkenas and C.L. Guerre, A Systems Analysis View of Longitudinal Flying Qualities, WADD 60-43, Wright Air Development Division, Jan., 1960.
 - (3) Smith, R.H., An Experimental Determination of the Limits of Pilot Controllability for Unstable, Oscillatory Second Order Systems, Thesis for degree of Aeronautical Engineer, Univ. of Cincinnati, 1961 (also appeared as a paper, On the Limits of Manual Control, IEEE Trans. on Human Factors in Electronics, Sept., 1963).

with the plant (aircraft) dynamics having the generic form (short-period approximation):

$$G_p = \frac{5(1+0.6s)\omega_n^2}{s(s^2 + 2\zeta\omega_n s + \omega_n^2)} \quad (1-14)$$

Hall used twenty combinations of ζ and ω_n to test pilot response. The plot for $-L_n = -G_p$ from (1-14) for $\zeta = 0.2$, $\omega_n = 1.57$ is shown in Fig. 1-10; this configuration was labelled C21 in the Hall report and is so labelled in Fig. 1-10.

Vellum (or transparent) overlays of figures such as Figs. 1-6 through 1-9 can be made and used to good advantage with system plots as presented in Fig. 1-10. Such overlays facilitate envisionment of whether the "difference vector" between $-L_n$ and Ω undergoes some type of rotation. In the case presented, a glance at Fig. 1-10 and the general trend of the Ω mappings for the frequencies given ($\omega = 0.5, 1.0, 2.0$ and 3.0), indicate that system stability is determined by the relation between the $-L_n$ plot and the mappings for $\omega = 2.0$ and $\omega = 3.0$. This situation results since the $-L_n$ plot leaves the Ω region of criticality for $\omega \approx 2$; this implies that a condition of instability may pertain. It is not yet clear how to take the difference vector between $-L_n(j2)$ and $\Omega(j2)$, i. e., conceivably, it could be connected from $-L_n(j2)$ to any point within or on $\Omega(j2)$; for instability, it must be demonstrated that a clockwise encirclement occurs as $\omega \rightarrow \infty$ irrespective of the selected orientation of this vector (all the Hall configurations were open-loop stable).

The time invariance hypothesis implies, however, that once a settled condition of operation has been adopted by the human controller, any "selected orientation" of the difference vector defines the sub-region in space to be used for the remainder of the stability analysis. For example, if the human operator assumes a set of parameters near point B' of Fig. 1-8 (corresponding to a neighborhood of $\max T_L$, $\max T_I$, $\max T_N$), there results, from Fig. 1-10, an encirclement by the difference vector and instability. If the point F' is assumed,

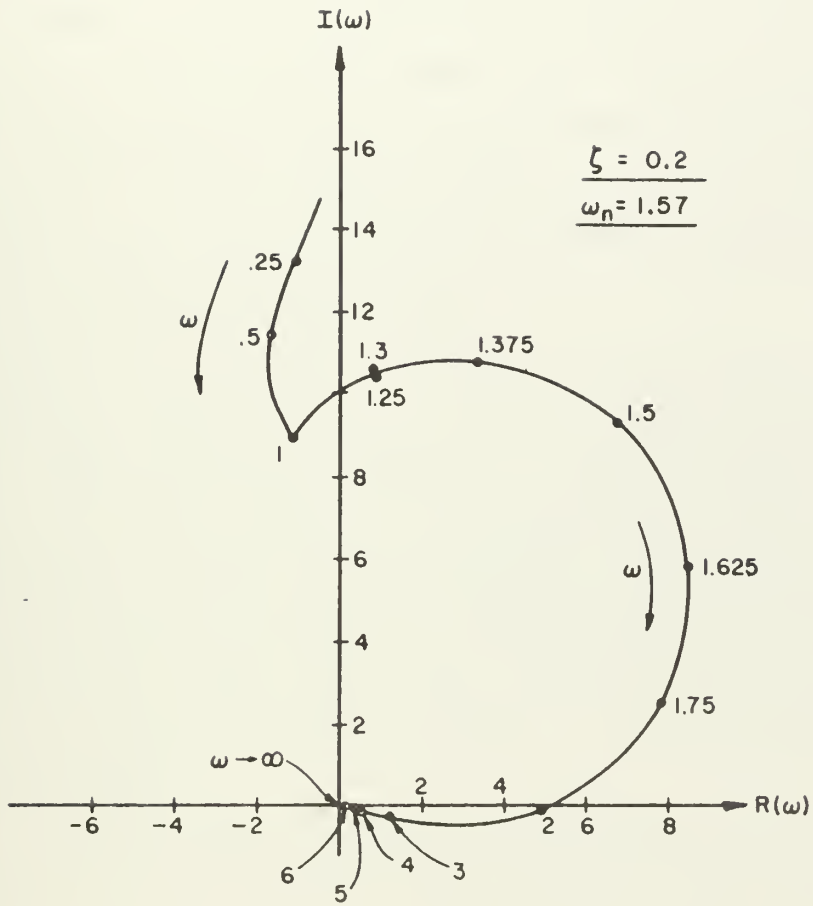


Fig.1-10: Plot of Configuration C21

the human can "hold out" longer before enclosure, but eventually an encirclement again occurs. Unfortunately, the prospects of being at F' are not encouraging since, although the point corresponds to $\max T_L$ (lead) which is obviously needed, the location F' also implies $T_N = 0$ which, if the physiological correlations are credible, corresponds to no neuromuscular lag. If there is any hope for system stabilization or controllability, the human operator must assume parameter values somewhere between arc $G'O'$ and arc $A'D'$ and in the vicinity of the origin of Fig. 1-8, also Fig. 1-5 (detail). Particularly, is this the case when the human time-delay is imposed upon the system. Thus, the characteristic of C21 projecting beyond the Ω region is further aggravated by the introduction of the time lag, i.e., the point $-L_n(j2)$ located at approximately 5.0 on the real axis rotates, with the distance from the origin to $5 + j0$ as radius arm, a number of degrees proportional to the time delay.

Since C21 represents a set of dynamics for which empirical evidence of human dynamic response exists, it is possible to check the inferences above. A "best fit" human operator description was calculated by Hall and a plot of this function can be made. In Section 5.3 it is shown that the human controller did, indeed, assume those characteristics previously implied by the analysis; furthermore, it is obvious that his capacity is being strained severely just to maintain control of the C21 system. That is to say, few points in the variable mapping space Ω are not enclosed by the plot of Fig. 1-10, particularly after introduction of time-delay effects.

It is interesting that the "subjective evaluation" of this system by two experienced Navy pilots was a rating of "unacceptable-quite unsuitable." Good correlation between the analytical predictions and the pilot ratings also occur for other sets of dynamics (Chapter Five). As implied already, correspondences between boundary lines of the mapping and the physical capacity of a human operator seem to be possible.

1.6 "Best " Performance and "Best-Overall" Compensation of Varying-Parameter Plants

As in all engineering design, several performance measures have been formulated for man-machine systems to provide a so-called figure of merit or index of quality; typically, deficiencies have been found to exist with all these scalar indices. The difficulty in attempting to quantify the performance of a complex system by a single number is that satisfaction of a requirement on this unique scalar may still result in a system with other unsatisfactory characteristics, irrespective of formulation and weighting techniques. A "vector" performance index might offer significant advantages. Such an approach to the performance problem is pursued in this thesis. Hopefully, it may be suggestive of a standard technique, or algorithm, for handling far more severely complex situations than can be exhibited here. Also, one feels intuitively that design compromise on an adequate compensator should be possible so as to provide "best" performance even over some range of plant-parameter variation. A simplified example of a method for satisfying such demands is given below.

The goal is a "best-overall" compensator for any given system with a human controller, Fig. 1-11. The approach may be used for a plant with variable parameters; the solution to be presented, however, depends intimately upon an initial stability study conducted in accordance with the procedure previously outlined. Thus, in Fig. 1-11, the result desired is values for K_c , α and β which provide good performance for all possible values of ζ . This portion of the thesis departs from stochastic-input-signal considerations and studies deterministic-input-signal response. Thus, performance may be measured by the classical indices t_r , ϕ_s and t_s (response rise-time, overshoot and settling-time resulting from a unit step input); these follow exactly from some arbitrary definition and may be regarded as a vector quantity in three space. As a result of a stability analysis as outlined above, approximate values for K_c , α , and β

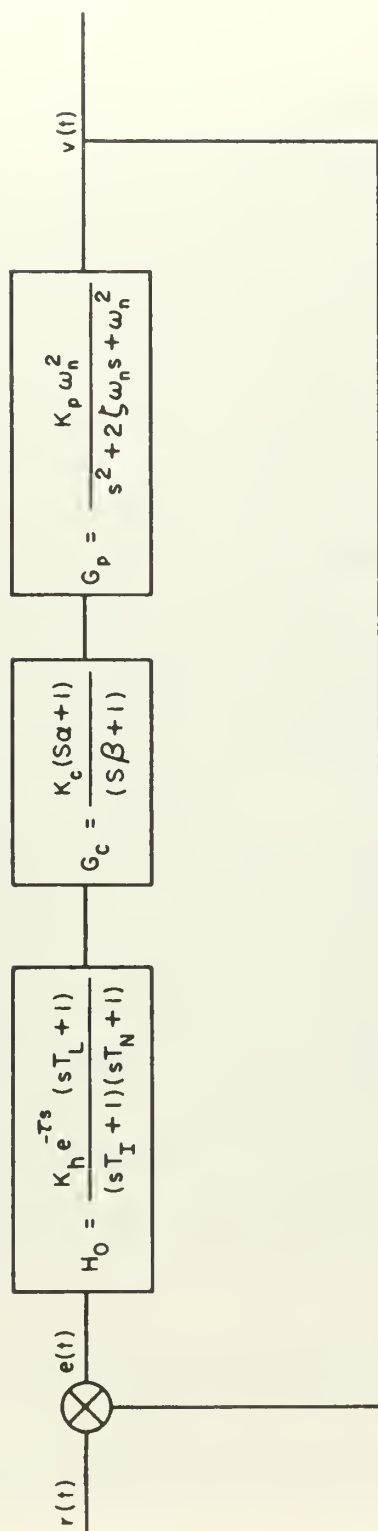


Fig.1-11 : Human Control with "Best" Compensation

conducive to stability are determined. In addition, the designer attempts to use the freedom in selecting α and β (for this imposed class of compensator) so that the plant's Nyquist curve passes through that part of the Ω region corresponding to reduced task requirements on the human operator; i.e., one establishes as large a "margin of human capacity" as possible. With the ranges of α and β that provide acceptable stability available, nominal values for the human controller can be read from the plot using the time-invariance assumption. Simultaneously, these values must retain the "high-frequency" end of the compensated Nyquist contour in a location that continues to satisfy the stability criterion, as indicated by the corresponding Ω region.* Thus, the triple (T_L, T_I, T_N) must have value such that it lies to the left of the plant's Nyquist curve (traversed in the direction of increasing ω) and still be within the range of human adaptivity. Also, T_L, T_I and T_N should be as small as possible (the human prefers to operate as a simple amplifier), particularly T_L , since this parameter is almost a direct indication of "comfortable" adaptation.

The above analysis, conducted for a nominal ζ , is repeated for several values covering the range of ζ . The greatest common range of α and β , obtained from the various ranges resulting from the repetitive analysis, then provide satisfactory stability over the domain of varying ζ . From these final ranges of acceptable α and β , the selection is made of those values of α and β which produce the best performance. There is no difficulty in determining t_r, o_s, t_s for any given set of K_c, α and β ; the problem is how to select subsequent trials of α and β from the final range established, so that each iteration is at least as good as the last.

The function

$$f = (t_r - T_r)^2 + (o_s - O_s)^2 + (t_s - T_s)^2 \quad (1-15)$$

* This statement assumes the human operator uses an equivalent performance-criterion.

is formed, where

$$0 < t_r \leq T_r$$

$$0 < o_s \leq O_s$$

$$0 < t_s \leq T_s$$

That is, T_r , O_s and T_s are the maximum acceptable bounds on t_r , o_s and t_s . The function is an artifice created specifically to facilitate finding some point on the surface of acceptable performance; at $f=0$, the specifications are met identically at their upper bounds. A minimization process conducted on the function (which contains time-domain parameters) produces useful information regarding acceptable K_c , α and β (which represent frequency-domain parameters). Appropriate reductions in T_r , O_s and T_s at timely points in the search procedure provide resolution of corresponding boundaries of acceptable-performance in both (t_r, o_s, t_s) and (K_c, α, β) space.

A very straightforward search technique is used to find a compensator satisfying the requirements.⁽¹⁾ This method of "direct search" with "pattern strategy" is especially suitable to minimization of a sum of squares, but its overriding virtues are its straightforward simplicity and compatibility with the electronic computer. "Direct search" means a step-by-step examination of trial solutions and a method of comparing each to determine which is "best" in some sense. Additionally, a method is afforded for choosing the next trial point. This last aspect forms the strategy part of the search procedure; a "pattern" approach is used and from this is derived the strategy's name.

The simplified example used to demonstrate this search technique is

⁽¹⁾ Hooke, R. and T.A. Jeeves, Direct Search Solution of Numerical and Statistical Problems, J. Assoc. Comp. Mach., April, 1961.

shown in Fig. 1-12; the human-operator description has been set at unity since its presence, while having no effect upon the search procedure, complicates the explanation. It is assumed that the ranges of α and β have been established which offer good stability; specific values within that range are now desired which meet the step-response requirements on rise time, overshoot and settling time. For economy of notation K_c is redesignated K . The initial "base point" of the search (K_1, α_1, β_1) is chosen arbitrarily or the designer may exercise a "hunch". Again, the function, f , (1-15) is formed; the point (K_2, α_2, β_2) is said to be "better" than (K_1, α_1, β_1) if

$$f(K_2, \alpha_2, \beta_2) < f(K_1, \alpha_1, \beta_1) \quad (1-16)$$

By $f(K_1, \alpha_1, \beta_1)$ is meant the value of (1-15) found from the substitution of the t_r , o_s , and t_s associated with the system step-response when the compensator has gain and time constants, K_1 , α_1 , and β_1 , respectively. If (1-16) holds, a "nothing succeeds like success" philosophy dictates a continuation of the search in the same "direction"; therefore, exploration is continued by perturbing the independent variables with increments whose proportions are based on the frequency of success achieved along each dimension, K, α, β , in the recent past.

In the system of Fig. 1-12, the specifications are chosen to be

$$T_r = 2 \text{ seconds}, \quad O_s = 15 \text{ percent}, \quad T_s = 4 \text{ seconds} \quad (1-17)$$

also, the plant variations possible are given to be

$$0.01 \leq \zeta \leq 1.0 \quad (1-18)$$

Exploration is conducted using step sizes as follows:

$$\text{For } K_c : \text{ take step } \delta_1 = 0.2 \quad (1-19a)$$

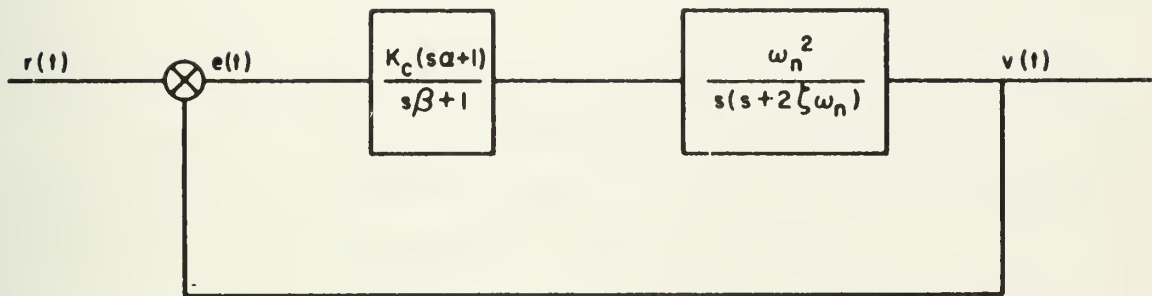


Fig. 1-12 : Compensation of a Varying-Parameter System

$$\text{For } \alpha : \text{ take step } \delta_2 = 0.1 \quad (1-19b)$$

$$\text{For } \beta : \text{ take step } \delta_3 = 0.05 \quad (1-19c)$$

The initial base point is chosen as

$$K_1 = 1 \quad (1-20a)$$

$$\alpha_1 = 2 \quad (1-20b)$$

$$\beta_1 = 0.5 \quad (1-20c)$$

In (1-19a, b, c) the subscripts 1, 2, 3 are assigned to represent step increments along the K_c , α and β dimensional axes, respectively. Usually, the problem and design experience suggest step sizes for each independent variable. During the termination procedure, the nominal step sizes are reduced until impractical as a consequence of solution indifference or equipment limitations.

The search is started at the varying plant's "sluggish end" ($\zeta = 1.0$). The initial base point for the search is called, \bar{b}_1 , (1-20a, b, c); the horizontal bar signifies the vector (multi-dimensional) nature of the point, i.e.,

$$\bar{b}_1 = (1, 2, 0.5) \quad (1-21)$$

The step response of the system with a compensator which possesses such characteristics is

$$(t_r, o_s, t_s) = (4.8, 0, 9) \quad (1-22)$$

From (1-15) and (1-17),

$$f(1, 2, 0.5) = [4.8-2]^2 + [0-15]^2 + [9-4]^2 = 257.81 \quad (1-23)$$

This value for f must now be compared with the value at the next exploration point. The next observation is performed at the point $\bar{b}_1 + \bar{\delta}_1$ where

$$\bar{b}_1 + \bar{\delta}_1 = (1.2, 2, 0.5) \quad (1-24)$$

The response is found to have performance measures

$$(t_r, o_s, t_s) = (3.5, 0, 6.8) \quad (1-25)$$

Therefore

$$f(1.2, 2, 0.5) = [3.5-2]^2 + [0-15]^2 + [6.8-4]^2 = 235.05 \quad (1-26)$$

Thus,

$$f(\bar{b}_1 + \bar{\delta}_1) < f(\bar{b}_1) \quad (1-27)$$

and $\bar{b}_1 + \bar{\delta}_1$ becomes the new temporary base point, \bar{t}_{11} , from which exploratory moves along the α axis are conducted.

The process is continued in this way, the best of three exploratory points (only two are shown above, generally, f is also evaluated at $\bar{b}_1 - \bar{\delta}_1$, also) always leading to a new temporary base point. The procedure is performed along each of the axes (K_c, α, β) and yields a new base point, \bar{b}_2 , after the completion of such a search cycle. With a successful "direction" of search established ($\bar{b}_2 - \bar{b}_1$), the next temporary base point used for exploration purposes is found by doubling the magnitude $\bar{b}_2 - \bar{b}_1$ and proceeding in the same direction. This technique forms the "pattern strategy" portion of the search procedure since future points of exploration are determined as a function of the frequency of success achieved along each dimension in the perturbation steps such as (1-24). Great strides can be made in pattern moves after very few iterations. This example was solved, effectively, after six iterations; the geometric aspects of the first two iterations are shown in Fig. 1-13. It is found that the triple

$$(K_c, \alpha, \theta) = (5, 1.0, 0.05) \quad (1-28)$$

satisfies the requirements

$$T_r = 1.0, \theta_s = 10, T_s = 2 \quad (1-29)$$

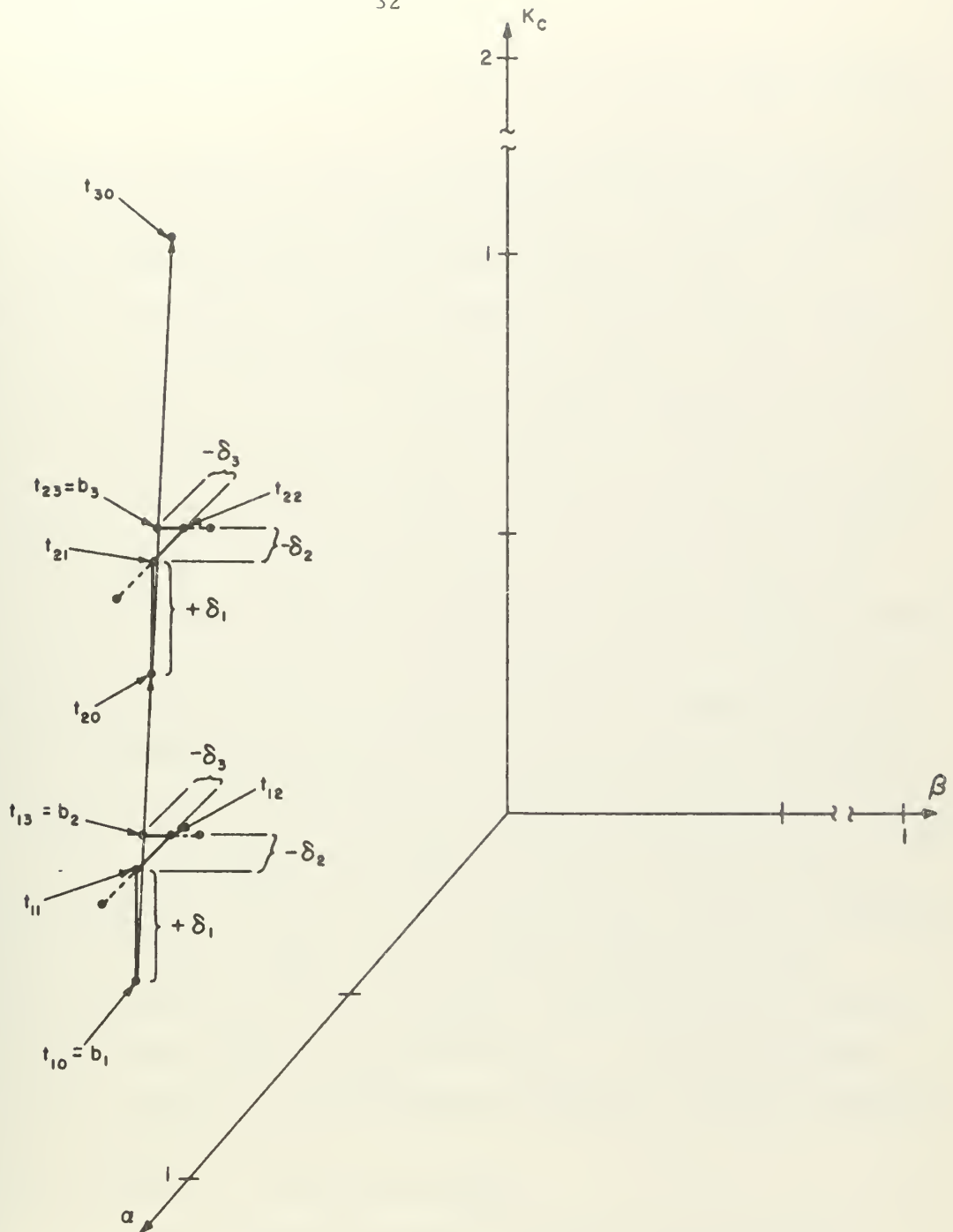


Fig. 1-13 : "Direct Search" with "Pattern Strategy"

Only the first two iterations are shown

over the range of $0.01 \leq \zeta \leq 1.0$; this is superior performance to that originally specified.

Empirical justification of the above approach (with a human operator in the loop) can be offered only after additional research is conducted for deterministic inputs driving controlled elements other than pure gains.* It is recognized that the bulk of the evidence justifying the form of human-operator transfer function used in this study has been accumulated for systems driven by stochastic input signals. There is no reason to suppose significant changes in this form for deterministic inputs; the initial range selections for K_h, τ, T_L, T_I and T_N were quite broad, but modifications, if needed, are easily incorporated.

Other search techniques might also be used to identify those values of K_c, α and β yielding "best-overall" compensation. Any method employed, however, must be suitable to finding the optimum of functions which are unknown to the observer; i.e., the analytic form of the function is unknown or too complicated to be used. This was the case in the previous example; the direct functional relationship between (K_c, α, β) and (t_r, o_s, t_s) is not easily specified. First, the transient solution for given K_c, α, β must be obtained by performing an inverse Laplace transformation. From the particular transient response resulting the values of t_r, o_s, t_s can then be determined. No direct relation between the time-domain performance-indices t_r, o_s, t_s and the frequency-domain compensator parameters K_c, α, β can be written. The end result is that experimental optimization procedures become most attractive. Wilde⁽¹⁾ describes several search techniques that could be used alternatively, e.g., Method of Steepest Ascent, Contour Tangents, Parallel Tangents or Method of Rotating Coordinates.

* Such results could not be located by the author.

(1) Wilde, D.J., "Optimum Seeking Methods," Prentice-Hall, Englewood, Cliffs, New Jersey, 1964.

1.7 Summary

Chapter One provides an overview of the main points considered in this study. The goals, accomplishments and limitations of the investigation are sketched briefly and a concise description of several of the major techniques used is presented. Further detailed discussion may be found in the main body of the thesis.

CHAPTER TWO

THE HUMAN OPERATOR IN CONTROL SYSTEMS2.1 Introduction

The possibility of describing human operator "behavior" in terms of mathematical relationships began in earnest sometime during World War II. Confronted with the problem of designing complex systems such as gun turrets capable of tracking fast and maneuverable aircraft, engineers found themselves needing considerably more quantitative information about the human operator's transfer characteristics than was previously required in weapons design. Whole teams of professionals in mathematics, engineering, psychology and physiology set to work to determine more adequately the bounds of human performance and the capabilities that could be expected.⁽¹⁾ In addition to Tustin's work, efforts in the United States were being conducted along similar lines by Phillips; some mention of this is made in the now classic text of which he is a co-author.⁽²⁾ The interested reader may find a good historical sketch in Sheridan's paper⁽³⁾ and the background interplay between the engineering and psychological philosophies given by McRuer.⁽⁴⁾ The resurgence of interest in the human controller probably may be marked by Russell's paper⁽⁵⁾ and the definitive work by

-
- (1) Tustin, A., An Investigation of the Operator's Response in Manual Control of a Power Driven Gun; C.S. Memorandum No. 169, Metropolitan-Vickers Electrical Co. Ltd., Attercliffe Common Works, Sheffield, England, Aug., 1944.
 - (2) James, H. M., N. B. Nichols and R. S. Phillips, "Theory of Servomechanisms," McGraw-Hill, 1947.
 - (3) Sheridan, T. B., The Human Operator in Control Instrumentation, "Progress in Control Engineering", Vol. 1, Academic Press, New York, 1962.
 - (4) McRuer, D. T. and E. S. Krendel, Dynamic Response of Human Operators, WADC Tech. Rpt. 56-524, Oct., 1957.
 - (5) Russell, L., Characteristics of the Human as a Linear Servo Element, (M. S. Thesis, M. I. T., May, 1951).

Elkind.⁽¹⁾⁽²⁾ Military high performance aircraft, space capsules, deep submergence vehicles and other systems whose dynamics are under the control or partial control of a human operator have served to intensify the interest in the human as a controlling element; additionally, the need for greater understanding of the man-machine interface has spurred research in this area.⁽³⁾⁽⁴⁾

Almost all the human operator studies conducted to date, and this thesis as well, address themselves to the problem of a human controller in a tracking situation; in particular, these investigations have been done in the "compensatory" mode, i. e., with the controller operating off the system error, Fig. 2-1. This topology lends itself most readily to analysis techniques; in the face of considerable ignorance about the human mechanism and the fantastic description-defying complexity of that mechanism some simplifying assumptions are justifiable. Particularly is this the case when the simplifying assumptions still permit useful conclusions to be derived which succeed in predicting empirical results.

Several characteristics peculiar to human-operator control are not treated in this study. Psychological aspects of human behavior (that is, stress, competitive situations, rewards, other similar task influences and their effects upon tracking) are not considered as such. Physiological loading influences (such as fatigue or alcohol) also have not been incorporated. This is not to discount the importance of these factors upon perceptual-motor behavior and their resultant effect on any mathematical model of the tracking situation; again, as many superfluous variables as possible must be eliminated if useful analytical models

-
- (1) Tracking Response Characteristics of the Human Operator:, Human Factors Operations Research Lab. Rpt. No 40, PRP-2, Sept., 1953.
 - (2) Elkind, J. I., Characteristics of Simple Manual Control Systems, (Ph. D. Thesis, M. I. T., Apr., 1956) also issued as M. I. T. Lincoln Lab. Report, No. 111.
 - (3) Hall, I. A. M., Effects of Controlled Element on the Human Pilot, Princeton Univ., WADC Tech. Dept. No. 57-509, Aug., 1958.
 - (4) Bekey, G. A., An Investigation of Sampled Data Models of the Human Operator in a Control System, (Ph. D. Thesis UCLA, Jan., 1963) also issued as ASD-TDR-62-36 WADC, May, 1962.

of the human controller are to be formed. The completely successful explanation of human response in this situation of reduced complexity (a milestone still not yet attained) places the future investigator closer to the possibility of including these additional parameters in a more accurate description of general human behavior.

There have been recent studies attempting a time-variant description of humans⁽¹⁾⁽²⁾⁽³⁾ and some experimental work relating to the adaptation process.⁽⁴⁾ Almost no work has been published on man as a monitor* although the problem is apparently under active study. Much of the published material on the monitoring situation has been essaylike⁽⁵⁾, although there has been some hardware generated by the Federal Aviation Agency as part of the supersonic transport program. This system, known as "Force Wheel Steering", allows the pilot to override the automatic elements; flight tests have shown considerable promise.⁽⁶⁾ A recent paper on monitoring⁽⁷⁾ reports a tentative theory to explain some experimental results but the theory is quite preliminary and, although some model fitting is attempted, nothing to compare with the demonstrated validity of the quasi-linear transfer function is produced.

* Operating as a monitor, the human controller observes the character of the error signal, but instead of remaining in the loop and processing the error signal continuously, he enters only as frequently as he believes he is required to augment normally "automatic" control.

- (1) Sheridan, T. B., Studies of Adaptive Characteristics of the Human Controller, Air Force Systems Command ESD-TDR 62-351, Dec., 1962.
- (2) Elkind, J. I., and D. M. Green, Measurement of Time-Varying and Non-Linear Dynamic Characteristics of Human Pilots, Aeronautical Systems Division, Tech. Rpt. No. 61-225, Dec., 1961.
- (3) Elkind, J. I., et al, Evaluation of a Technique for Determining Time-Invariant and Time-Variant Dynamic Characteristics of Human Pilots, NASA TND-1897, May, 1963.
- (4) Young, Green, Elkind and Kelly, Adaptive Dynamic Response Characteristics of the Human Operator in Simple Manual Control, IEEE Transactions on Human Factors in Electronics, Sept., 1964.
- (5) Hickey, A. E. and W. C. Blair, Man as a Monitor, Human Factors, Sept., 1958.
- (6) Control-Display Pilot Factors Program, Instrument Evaluation Project Report 63-1, USAF Instrument Pilot School, Randolph AFB, Texas, Dec., 1963.
- (7) Senders, J. W., The Human Operator as a Monitor and Controller of Multi-degree of Freedom Systems, IEEE Transactions on Human Factors in Electronics, Sept., 1964.

The following definition of quasi-linearization appears in Elkind.⁽¹⁾

"Quasi-linearization is the process of representing the characteristics of a non-linear device by linear transfer functions whose parameters depend upon the environment of the device. These transfer functions are called quasi-linear transfer functions because they behave linearly when the set environmental parameters remain fixed and nonlinearly when these parameters are allowed to change."

"Quasi-linearization is particularly suited to the kinds of nonlinearities observed in the human operator. His most significant departures from linearity occur when the input statistics or system dynamics change. Since its parameters depend upon the input and the system, a quasi-linear transfer function can handle this kind of nonlinearity."

It is interesting to quote the following comment from a recent NASA report. "The potentials of quasi-linear models have not been fully investigated and further studies along this line are warranted."⁽²⁾

In this paper a technique is developed to display graphically a range of human-operator capability which has been derived from a variety of such quasi-linearization studies; these plots are then used, in conjunction with standard Nyquist contours which represent the controlled elements, to predict human control capacity. It is shown that these predictions parallel previously published experimental results. Following the stability studies, there is a chapter devoted to performance considerations, Chapter Six. The classical measures of step input response characteristics, rise time, overshoot and settling time are used to demonstrate that a "vector" performance index can be formed and satisfied. It is shown that the required compensator design for the satisfaction of such specifications can be determined using computer search methods in combination with an "artificial" function which, when minimized to equal zero, identically meets the stipulated response requirements.

(1) Elkind, J. I., Characteristics of Simple Manual Control Systems, (Ph. D. Thesis, M. I. T., Apr., 1956) also issued as M. I. T. Lincoln Lab. Report, No. 111.

(2) Summers, L. G., K. Ziedman, A Study of Manual Control Methodology with Annotated Bibliography, NASA Contractor Report CR-125, Nov., 1964.

2.2 Controllability and Adaptation

The utility of the quasi-linear transfer function approach to engineering system design has been established many times, over recent years; even in the face of its many acknowledged limitations it represents a technique which provides useful results. This thesis uses as its starting point the quasi-linear "best-fit" functions and the empirical results that have been published in the literature. A fresh look at the analytical tools extant and some novel manipulation of system transfer functions afford a new method for incorporating the entire adaptive range of the human operator and his pure time delay in a stability analysis. To date servo analysis techniques used in man-machine studies have utilized a nominal transfer function where the specific values for the time constants are selected to reflect the gross form of the process, i. e., human-operator transfer functions for particular dynamic elements (plants or processes) have been experimentally determined and presented in a "table" which is then used by the systems designer. The table may be regarded as offering "adjustment rules" indicating how to "set" the parameters in the human-operator transfer function for varied control tasks and input power spectra. It should be noted that these studies normally employ simple transfer function approximations, usually of a bilinear form, for the pure time lag introduced by the human operator.

Some of the "inappropriate" and lightly justified aspects of the quasi-linear modeling techniques are subdued in this "new" theoretical approach to human-operator studies since a broad range of human-operator adaptation and control attributes are inherently included in the analysis technique. This broad range tends to reduce the sensitivity of analyses to inaccuracies in the nominal values for the human-operator parameters, which previously had to be specifically designated. Since no new experimental work is presented there is no attempt to put forth a new model or model matching technique; rather, the investigation should be regarded as providing another means of viewing the human operator in engineering terms with the physical justification of the theory obtained from previously published empirical results.

The mathematical model required to adequately describe the human operator for different tasks necessarily varies in correspondence with the human's actual

adaptive capabilities. To date, this variation has been characterized by selecting specific values for the quasi-linear description as a function of the particular task. The method offered in this thesis is sufficiently encompassing of the broad range a mathematical model of the human operator must take, that it compensates for some of the inherent rigidity this selection procedure introduces. The method accomplishes this compensation by its complete display of human-operator capacity. Simultaneously, the method provides both a means of analysis and a channel for developing better intuitive understanding of human-parameter effects upon a control system and system effects upon the human.

For man operated systems the most important questions to be posed are: "Can the human 'handle' the system?" and "What can be done to ease his control burdens?" A comprehensive analysis technique for certain man-machine systems is constructed by capitalizing on studies already conducted and results previously published in the literature. It is, perhaps, a first attempt at a theory for coming to grips not merely with the human controlled system, since this, indeed, has been tried before, but with that system under the added complication of the human-operator-adaptive range. Thus, to answer the question, "Can a human operator stabilize this system?" would be useful. The method to be offered will allow estimation of whether a human operator can stabilize any given system assuming he can adapt his "parameters" over certain ranges; in general, the investigator need not select specific values for these parameters. Insight into the relative human effort required to stabilize is also provided. No approximation to the pure delay introduced by all human operators need be made. The ultimate justification for the new approach is then provided by predicting whether the human operator can stabilize certain systems. The systems used were analytically equivalent to those reported on in the literature, along with their associated human operator evaluations; the published classifications were determined experimentally. The predictions parallel the empirical results remarkably well even for systems with open-loop instability. Qualitative information concerning handling qualities and engineering insight into how the systems may be modified to improve their stability and the human operator's satisfaction with their characteristics is also provided. The theory offers a convenient and illuminating display of the effects

upon the system of the various parameters which form the human-operator transfer function. Thus, it becomes easier to observe which ones (and their associated physiological characteristics, Section 2.5) are causing difficulty or limiting controllability.

2.3 Thesis Organization

The dissertation is effectively in two parts. The first (Chapters Two-Five) deals with system stability and controllability by the human operator and has been described qualitatively in the above paragraphs. The second part is concerned with system performance and introduces a type of "vector" performance measure which could have considerable usefulness in future human operator studies. This portion of the thesis departs from the stochastic input signal and studies the ramifications of deterministic forcing functions. Much experimental work in the psychological literature pertains to step input tracking and the techniques described in Chapter Six may have direct application there. Too often, however, the psychologist's reports are qualitative and include varied methodologies thus making cross-experimental comparisons with the engineer's work almost impossible. With all its difficulties the engineer's approach at least provides a consistent language. Chapter Six is offered not as a solution to the performance problem but as a stimulant to further study of performance measures. In conjunction with this "vector" type performance criterion, a "direct search" method is used to obtain a specific compensation of a given class which provides "good overall" performance in the presence of a plant (process) having variable dynamics. The results are applicable to automatic control systems, in general, and systems with a human operator, in particular.

Additionally, in this last chapter, the displays of human operator capacity, offered in the first part of the thesis are shown to have utility in the search for a solution to the performance problem; indeed, it is demonstrated in an example that these displays can be used as a tool in the design of compensators for "inanimate" systems.

2.4 The Quasi-Linear Model

In general, human-operator response characteristics are strong functions

of the input signal form. For signals such as single sine waves the operator tends to "synchronize" his response with the input; such situations where the human may exercise his capability for prediction have been called "precognitive". Precognitive response can be considered a form of open-loop control rather than tracking, although certain tracking tasks do permit this type of behavior. If both the operators corrective responses and the input are displayed separately, the form of tracking is called "pursuit"; the classic example of such tracking is aiming a rifle at a moving target through open sights. In "compensatory" tracking the operator cannot distinguish between the effects of his response actions and changes in the input, except under zero input conditions, i. e., the operator views an error display only. The analogous example for "compensatory" tracking is aiming a rifle at a moving target through a telescopic sight with a narrow field of vision. As mentioned above, this study restricts itself to the compensatory case.

A very brief derivation of how the quasi-linear human-operator transfer function is obtained, follows. From Fig. 2-1,

$$\dot{U}(j\omega) = \frac{H_o(j\omega)}{1 + H_o G_p(j\omega)} R(j\omega) + \frac{1}{1 + H_o G_p(j\omega)} N(j\omega)$$

where $N(j\omega)$ is the transform of $n(t)$, a term arbitrarily added at the operator's output to account for that portion of the output not described by the linear transfer characteristic; this term is called "the remnant."

Also,

$$E(j\omega) = \frac{1}{1 + H_o G_p(j\omega)} R(j\omega) - \frac{G_p(j\omega)}{1 + H_o G_p(j\omega)} N(j\omega)$$

The cross-spectra $\hat{\phi}_{ru}$ and $\hat{\phi}_{re}$ are found to be

$$\hat{\phi}_{ru}(j\omega) = \lim_{T \rightarrow \infty} \frac{R_T^*(j\omega) U_T(j\omega)}{2T} ; \quad \hat{\phi}_{re}(j\omega) = \lim_{T \rightarrow \infty} \frac{R_T^*(j\omega) E_T(j\omega)}{2T}$$

where the asterisk signifies the conjugate and the subscript T indicates the Fourier Transform has been taken on the associated truncated time function, e. g.,

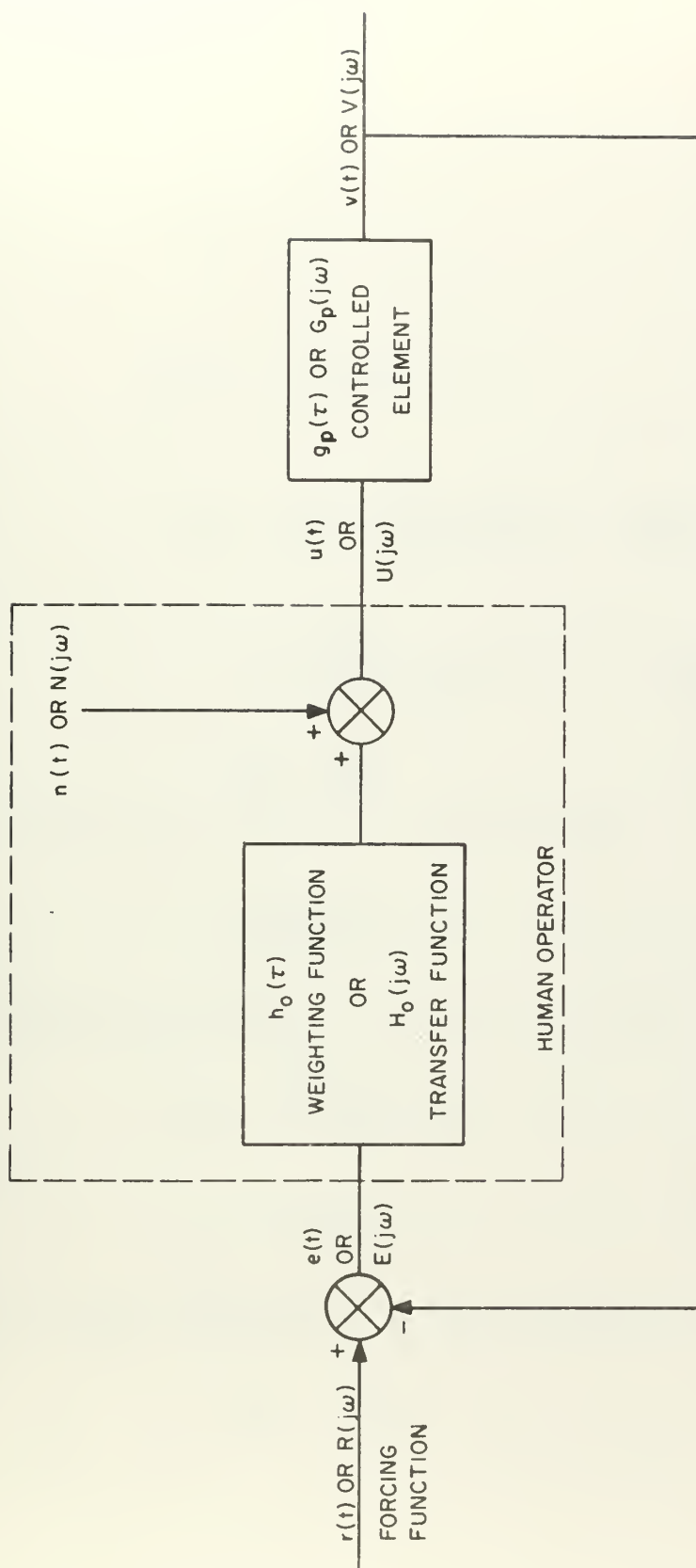


Fig. 2-1: The Human Controller in the Compensatory Mode

$$u_T(t) = \begin{cases} u(t), & |t| < T \\ 0, & |t| > T \end{cases}$$

and

$$U_T(j\omega) = \int_{-\infty}^{\infty} u_T(t) e^{-j\omega \tau} d\tau$$

Thus,

$$\begin{aligned} \Phi_{ru} &= \lim_{T \rightarrow \infty} \frac{1}{2T} \left[\left(\frac{H_o}{1 + H_o G_p} R + \frac{1}{1 + H_o G_p} N \right) R^* \right] \\ &= \frac{H_o}{1 + H_o G_p} \lim_{T \rightarrow \infty} \frac{1}{2T} [RR^*] + \frac{1}{1 + H_o G_p} \lim_{T \rightarrow \infty} \frac{1}{2T} [NR^*] \end{aligned}$$

the quantity

$$\lim_{T \rightarrow \infty} \frac{RR^*}{2T}$$

represents the input power spectrum Φ_{rr} and the quantity

$$\lim_{T \rightarrow \infty} \frac{NR^*}{2T} = 0$$

under the assumption that the signal and noise are uncorrelated. Therefore,

$$\Phi_{ru} = \frac{H_o}{1 + H_o G_p} \Phi_{rr}$$

and, similarly,

$$\Phi_{re} = \frac{1}{1 + H_o G_p} \Phi_{rr}$$

The amplitude and phase characteristics of the human operator model are then found from

$$H_o(j\omega) = \frac{\Phi_{ru}}{\Phi_{re}} \quad (2-1a)$$

i. e., the cross-spectrum between input and operator output divided by that between input and error yields a set of data in complex form which represents the human operator block of Fig. 2-1 (solid lines).

These data can be fitted closely by a linear analytic relation of the form:

$$H_o(j\omega) = \frac{K_h e^{-j\omega\tau} (1 + j\omega T_L)}{(1 + j\omega T_N)(1 + j\omega T_I)} \quad (2-1)$$

Jackson has shown that negligible increased accuracy results if (2-1) is made any more complicated.⁽¹⁾ Equation (2-1) provides a "best-linear-fit" in a minimum mean-square error sense^{*} assuming random inputs of a Gaussian type, i. e., normal amplitude distribution.⁽³⁾ The human operator adjusts his parameters to some unknown internally-generated criterion; comparisons with the mean-square-error criterion seem to support its validity as a substitute analytic index.⁽⁴⁾ The Gaussian assumption has proved to be reasonable and realistic in the light of the low-pass nature of most control systems, i. e., low-pass characteristics in feedback systems tend to make the amplitude distributions Gaussian (central limit theorem).⁽⁵⁾⁽⁶⁾ Additionally, the Gaussian signal lends itself to

* This is not so restrictive as might appear since any criterion which is a monotonically increasing, even error function is also minimized by the resulting linear fit.⁽²⁾

- (1) Jackson, A. S., Synthesis of a Linear Quasi-Transfer Function for the Operator in Man-Machine Systems, Wescon Convention Record, Sept., 1958.
- (2) Sherman, S., Non-Mean-Square Error Criteria, IRE Trans. on Information Theory, Sept., 1958.
- (3) Booton, R. C., The Analysis of Non-Linear Control Systems with Random Inputs, Proc. Symp. on Non-Linear Circuit Analysis, Polytechnic Institute of Brooklyn, N. Y., Apr., 1953.
- (4) Roig, R. W., A Comparison Between Human Operator and Optimum Linear Controller RMS-Error Performance, IRE Transactions on Human Factors in Electronics, Mar., 1962.
- (5) Gibson, J. E., "Nonlinear Automatic Control", McGraw-Hill, New York, 1963.
- (6) Arthur, G. R., A Note on the Approach of Narrow-Band Noise after a Non-Linear Device to a Normal Probability Density, J. Appl. Phys., Vol. 23, 1952.

experimental work; also, Booton has shown that "the best mean-square error approximation to any nonlinear devices having a Gaussian input will be realizable and is determined by (2-1a)".⁽¹⁾

From

$$\phi_{uu} = \lim_{T \rightarrow \infty} \left[\frac{U^* U}{2T} \right]$$

is found

$$\phi_{uu} = \left| \frac{H_o}{1 + H_o G_p} \right|^2 \phi_{rr} + \phi_{n_t n_t}$$

where

$$\phi_{n_t n_t} = \left| \frac{1}{1 + H_o G_p} \right|^2 \phi_{nn}$$

which may be considered the "closed-loop remnant" spectral density. Let

$$H = \frac{H_o}{1 + H_o G_p}$$

then

$$\phi_{uu} = |H|^2 \phi_{rr} + \phi_{n_t n_t} \quad (2-2)$$

Greatest confidence in the description (2-1) occurs when it represents the human operator "adequately", i. e., when $n(t)$, the remnant term, is small. A figure of merit termed "linear correlation" has been generated which relates the amount of linearly correlated operator output to total operator output. From (2-2) is found the fraction of the linearly correlated operator output power spectrum to total operator output power spectrum

$$l = \frac{|H|^2 \phi_{rr}}{\phi_{uu}} + \frac{\phi_{n_t n_t}}{\phi_{uu}}$$

(1) Elkind, J. I., Characteristics of Simple Manual Control Systems, (Ph. D. Thesis, M. I. T., 1956) also M. I. T. Lincoln Lab. Rpt. No. 111.

If the square of the linear correlation is written

$$\rho^2 = \frac{|H|^2_{rr}}{\Phi_{uu}}$$

then

$$\rho^2 = 1 - \frac{\sum n_t n_t}{\Phi_{uu}}$$

Of course, the closer ρ is to unity the greater the confidence in (2-1) as a description of human-operator response characteristics. Frequently, the linear correlation is good ($\rho \geq 0.8$), ⁽¹⁾ thereby justifying use of the quasi-linear model for low-frequency conditions. Further discussion of remnant characteristics and possible sources for the remnant term can be found in McRuer. ⁽²⁾

2.5 Adaptive Ranges of the Human Operator

From various published data in the literature ⁽³⁾⁽⁴⁾ the values assumed by K_h, τ, T_L, T_I and T_N were assembled and from the largest and smallest of these values ranges for each parameter were established. These data were associated with controlled elements (system dynamics) having forms of pure gain, lags, integrators and second-order systems. Even simulated approximations to human operator control of the short period longitudinal dynamics of an aircraft and the dynamics of an F-80A aircraft engaged in a tail chase have been studied. In general, the most complicated human-operator transfer functions that could be justified by these data have the form (2-1). Some double leads and lags were used in the linear fits by Hall ⁽⁴⁾ but these

-
- (1) Ashkenas, I. L., D. T. McRuer, A Theory of Handling Qualities Derived from Pilot-Vehicle System Considerations, Aerospace Engineering, Feb., 1962.
- (2) McRuer, D. T., and E. S. Krendel, Dynamic Response of Human Operators WADC Tech. Rpt. 56-524, Oct., 1957.
- (3) Ibid
- (4) Hall, I. A. M., Effects of Controlled Element on the Human Pilot, Princeton Univ., WADC Tech. Rpt. 57-509, Aug., 1958.

were isolated cases and they are of no great concern here. Again, the human operator adjusts his parameters to conform to each control task; the tasks themselves are changed by modifying the controlled element or the input power spectrum.

Since a considerable amount of work has been expended in obtaining these data and they are highly indicative of human-operator capabilities they should be sufficient to permit upper and lower bounds to be drawn upon the human adaptation process. It is not maintained that the ranges given below definitively bound human-operator capabilities, but they are suggestive of limits obtained for a variety of forcing functions and controlled elements. The question as to whether a human operator can freely adjust these parameters independently to optimize is not answered although some studies seem to imply he does.⁽¹⁾ The "adaptive behavior" of the human operator refers to responses to tasks AFTER the learning or training period has been accomplished. Inherent in the final adaptive-range potential is the assumption that the operator has gained some familiarization with the system. Studies are currently being conducted on the learning or adaptive process; some results have already been reported in the literature.⁽²⁾⁽³⁾

Thus, the human-operator transfer function with the range of variation indicated below should provide a very broad, useful set of results until further definition of bounds is empirically determined. The validity of the approach is not conditioned upon the absolute accuracy of these ranges which always can be modified when desired.

-
- (1) Roig, R. W., A Comparison Between Human Operator and Optimum Linear Controller RMS-Error Performance, IRE Transactions on Human Factors in Electronics, Mar., 1962.
 - (2) Young, L. R., D. M. Green, J. I. Elkind, and J. A. Kelly, Adaptive Dynamic Response Characteristics of the Human Operator in Simple Manual Control, IEEE Transactions on Human Factors in Electronics, Sept., 1964.
 - (3) Gibson, J. E., et al, Philosophy and State of the Art of Learning Control Systems, CISL, Purdue University, Nov., 1963.

The "best fit" (in the minimum mean-square sense) human-operator transfer function based upon Gaussian input spectral analysis takes the general form (2-1) ^{*},

$$H_o = \frac{K_h e^{-j\omega\tau} (1 + j\omega T_L)}{(1 + j\omega T_N)(1 + j\omega T_I)} \quad (2-1)$$

The ranges on the various parameters are

$$K_h : 0.6 - 250$$

$$\tau : 0.1 - 0.4$$

$$T_L : 0 - 5.3$$

$$T_I : 0.025 - 25$$

$$T_N : 0 - 0.7$$

where the units for K_h depend upon the units of the input to the human-operator block and the units for τ , T_L , T_I , and T_N are in seconds.

To date, little work has been done on the monitoring problem, and few of the techniques of control system analysis have been applied to either multi-axis or monitoring tasks involving the human operator. ⁽²⁾ The stability analyses presented below are also applicable to the monitoring situation. The analysis is conducted to establish stability conditions only for the situation where the human enters the system and remains as a controller. Further

^{*} Some work has been done on sampled data models of the human operator based upon the philosophy that human performance is of an intermittent nature rather than continuous. This concept is not used in this study; details are available in the literature. ⁽¹⁾ Bekey's work also contains a delightful critique on the psychological literature.

- (1) Bekey, G. A., An Investigation of Sampled Data Models of the Human Operator in a Control System, (Ph. D. Thesis UCLA, Jan., 1963) also issued as ASD-TDR-62-36 WADC, May, 1962.
- (2) Dander, V. A., Predicting Pilot Ratings of Multi-Axis Control Tasks from Single Axis Data, IEEE Trans. on Hum. Factors in Electronics, Sept., 1963.

work on periodic entrance and exit could be useful in application to time-sharing a human operator, i. e., the human is responsible for N systems which he may "trim up" on a "multiplexing" basis. As an amplification of the sense of the word "monitoring" as used in this study, the following quote is offered. ⁽¹⁾ "Force wheel steering operationally maintains the human pilot as an active control element with an automatic system for the purpose of having the human pilot monitor, supervise, and act as a redundant element for those contingencies which require pilot action... THE AUTOPILOT ACTS ON THE SAME GUIDANCE INFORMATION THAT IS PRESENTED TO THE PILOT; therefore, the pilot can easily monitor autopilot performance. For those approaches when the autopilot is operating satisfactorily, the pilot makes a minimum number of inputs with the Force Wheel. But the pilot can, upon assessment of the total situation, modify the autopilot performance using the same information as the autopilot by making normal control actions... pilot inputs can be made WITHOUT disconnecting the autopilot." A possible block diagram display of such a monitoring system is shown in Fig. 2-2. In order to incorporate the "hands off" mode in monitoring, i. e., when the human controller effectively eliminates himself from the loop as during the "off duty" period in true monitoring, the lower bound on K_h has been reduced to zero. To facilitate certain manipulations, and ease the algebra in some of the mapping theorems, the lower bound on T_I also is considered as zero; this in no way restricts the generality of the results. The broadened range does ensure that when the stability conditions are met, greater confidence in the physical realization of stability follows. On the other hand, the number of cases which are stable is reduced; the modification required on T_I is not too significant however, and is justified by the simplification provided.

There has been an attempt to associate some of the parameters in the human-operator transfer function with physiological characteristics; the consensus is generally as follows:

(1) USAF Instrument Pilot School, Randolph AFB, Texas, Control Display Pilot Factors Program, Instrument Evaluation Project Report 63-1, Dec., 1963.

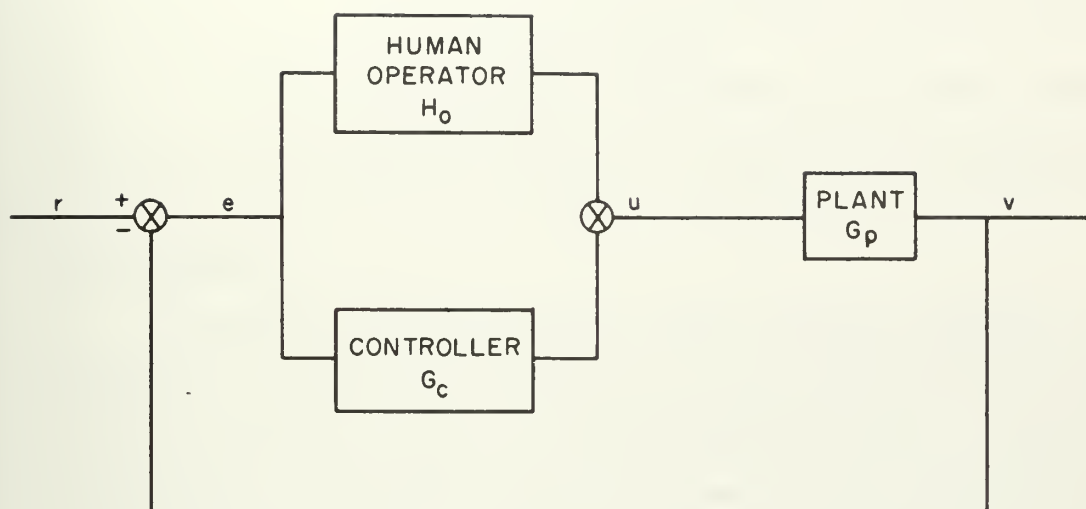


Fig. 2-2: The Human Controller as a Monitor

- K_h : human-controller gain; normally adjusted for overall system stability and low-frequency performance
- τ : pure delay (characterizes human reaction time)
- T_L, T_I : equalization; these provide the human operator's adjustable compensation similar to lead or lag networks in "inanimate" systems
- T_N : neuromuscular lag; appears to be partially adjustable for the task

This study does not enter the debate on the validity of these assignments.

2.6 Summary

Normally, the human-operator transfer function is specified as a function of the controlled-element form and type of input power spectrum. Thus, any single description for the human operator really is valid only for a set of limited conditions. This study demonstrates the human-operator transfer function approach may be generalized by using a variety of published experimental data to form "capability bounds" upon the transfer function parameters. With such ranges defined, the human-operator capacity for adaptation can be displayed and used in conjunction with controlled-element system dynamics to determine certain "man-machine" closed-loop stability characteristics. This technique, to some extent, depends upon the assumption that the human operator uses as a basic criterion the maintenance of system stability; this assumption has been verified as a reasonable one after many experiments. A man-monitored system configuration also appears to be susceptible to such stability analysis. The difficulty experienced by a human operator in attempting to maintain control of the system is implied by the "margin of human capacity" remaining; this capacity appears graphically as a part of the analysis technique.

CHAPTER THREE

HUMAN OPERATOR ADAPTIVE RANGES AND STABILITY THEORY3.1 Introduction

In this chapter the system transfer function mathematics are presented and the relations are manipulated so the adaptive properties of the human operator are isolated. The concept of a nominal transfer function is introduced to facilitate the separation of the closed-loop-system function denominator into a fixed and a varying part. Finally, a brief discussion of stability interpretation for this modified Nyquist display is given.

Many of the concepts in this study may be found in different form throughout the engineering literature. Several such published papers led to considering the possibility of a study on the effects of variations in the parameters of a transfer function the FORM of which was known.⁽¹⁾⁽²⁾⁽³⁾⁽⁴⁾⁽⁵⁾ Also, techniques similar in certain respects to those presented here are offered in Horowitz's text.⁽⁶⁾ Of considerable aid in contemplating the theoretical ramifications of a pure time delay was Choksy's work⁽⁷⁾. The utility of the method described lies in its inclu-

- (1) Blecher, F. H., Design Principles for Multiple Loop Transistor Feedback Amplifiers, Bell Labs Memo. 57-2172-6, Nov. 8, 1957 also appears in Proceedings of the National Electronics Conference, Vol. XIII, Oct. 7-9, 1959. (The stability criterion in this paper is due originally to M. L. Curtis).
- (2) Myers, B. R., A Useful Extension of the Nyquist Criterion to the Stability Analysis of Multiloop Feedback Amplifiers, Proc. 4th Midwest Symposium on Circuit Theory, Marquette University, Dec., 1959.
- (3) Jones, P., Stability of Feedback Systems Using Dual Nyquist Diagram, Trans. IRE Prof. Group on Circuit Theory, Mar., 1954.
- (4) Satche, M., Discussion, J. Appl. Mech., p. 419, Dec., 1949.
- (5) Stewart, R. M., A Simple Graphical Method for Constructing Families of Nyquist Diagrams, J. Aero Sciences, p. 767, Nove., 1951.
- (6) Horowitz, I. M., "Synthesis of Feedback Systems," Academic Press, N. Y., 1963.
- (7) Choksy, N. H., Time Lag Systems, "Progress in Control Engineering", Vol. 1, Academic Press, N. Y., 1962.

sion of a range on the parameters of interest so a specific expression of a transfer function need not be selected. The case of parameter variation in a transfer function of given form is precisely the state of affairs in certain man-machine systems where the human operator provides an adaptive transmission link.

3.2 System Transfer Function Manipulation

The closed-loop system function for Fig. 2-2 is found to be

$$\frac{V}{R} = \frac{H_o G_p + G_c G_p}{1 + H_o G_p + G_c G_p} \quad (3-1)$$

It is noteworthy that in solving the stability problem for this system several other configurations are handled simultaneously. For example, removing G_c in Fig. 2-2 and placing it in feedback around G_p leads to the same characteristic equation as in (3-1). The stability of (3-1) depends upon whether any of the zeros of its denominator lie in the right-half plane. For stability, there must be no zeros of

$$1 + L = 1 + H_o G_p + G_c G_p \quad (3-2)$$

in the right-half plane*.

Equation (3-2) can be written

$$1 + L = (1 + G_c G_p) \left(1 + \frac{H_o G_p}{1 + G_c G_p} \right)$$

or

$$1 + L = (1 + L_p)(1 + L_h)$$

where L_p is the open-loop transfer function corresponding to the compensator-plant and L_h is the open-loop transfer function corresponding to the human controller-plant path where the plant is somewhat more complicated than G_p and

* That is, no roots of $1 + H_o G_p + G_c G_p = 0$ in the right-half plane.

is represented by

$$\frac{G_p}{1 + G_c G_p}$$

Of course, $1 + L_p$ presents no serious problem in stability analysis; the techniques for stability checks and performance studies in this situation are well known since the structure is **FIXED PARAMETER**.

The term containing the **VARYING PARAMETERS** is given by, $1 + L_h$ where

$$L_h = \frac{H_o G_p}{1 + G_c G_p} = H_o G'_p$$

that is,

$$G'_p = \frac{G_p}{1 + G_c G_p}$$

The symbol, H_o , represents that part of the L_h expression which has variable parameters and G'_p that part with parameters nominally fixed.*

Since;

$$1 + L_h = 1 + H_o G'_p \quad (3-3)$$

this can be written as

$$= 1 + G'_p H_{o_n} \left[\frac{H_o}{H_{o_n}} \right]$$

where H_{o_n} represents some nominal value of H_o which is obtained by picking any set of values of the varying parameters in the given range of their variation.

Now, the zeros of

$$1 + G'_p H_{o_n} \left[\frac{H_o}{H_{o_n}} \right] \quad (3-4)$$

are the zeros of

$$\frac{H_{o_n}}{H_o} + G'_p H_{o_n} = \frac{H_{o_n}}{H_o} + L_n \quad (3-5)$$

* The examples in this study are for $G_c = 0$ in order to simplify the presentation, permit an intuitive grasp of the method, and allow comparisons with experimental results reported in the literature, since few empirical data have been reported for the monitoring configuration. No loss in generality occurs; the condition of an L_h with right-half plane poles, a situation representing the worst effect of a non-zero G_c , will be discussed in an example.

where

$$L_n = G'_p H_{o_n} \quad (3-6)$$

is the nominal loop-transmission with FIXED PARAMETERS, Fig. 3-1. The stability problem is now one of shaping the function with non-varying parameters L_n such that in the face of the parameter variation in $\frac{H_{o_n}}{H_o}$ the zeros of $\frac{H_{o_n}}{H_o} + L_n$ are all in the left-half plane; the criteria ensuring stability will be discussed below.

3.3. The Nominal Function

Of course, H_{o_n} should be chosen to facilitate the analysis. Equation (2-1), representing the human operator is

$$H_o = \frac{K_h e^{-s\tau} (1+sT_L)}{(1+sT_N)(1+sT_I)} \quad , \quad s = j\omega \quad (3-7)$$

where the final ranges considered in this study are

$$\begin{aligned} 0 &\leq K_h \leq 250 \\ 0.1 &\leq \tau \leq 0.4 \\ 0 &\leq T_L \leq 5.3 \\ 0 &\leq T_I \leq 25 \\ 0 &\leq T_N \leq 0.7 \end{aligned}$$

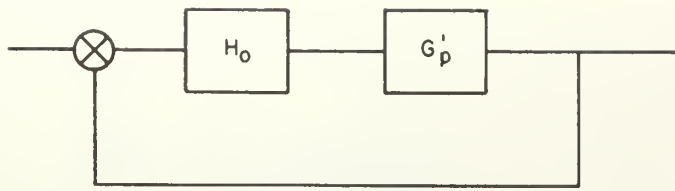
As an example, the nominal expression for H_{o_n} is taken as

$$H_{o_n} = e^{-.1s} \quad (3-8)$$

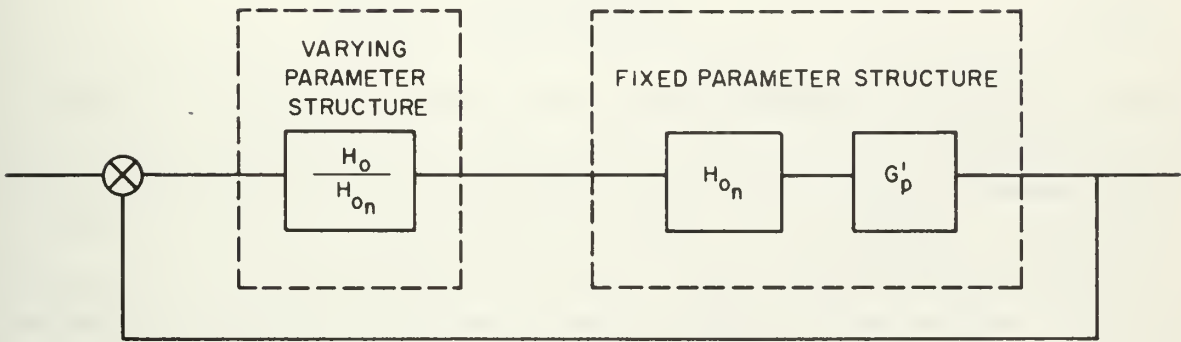
where it is clear that (3-8) is (3-7) with

$$K_h = 1, \tau = 0.1, T_L = T_I = T_N = 0$$

Thus,



(a)



(b)

Fig. 3-1 : Separation of Fixed and Varying-Parameter Elements

(a) Original System, (b) System after introduction of nominal function, H_{o_n} .

$$\frac{H_{on}}{H_o} = \frac{e^{-.1s}}{K_h e^{-.1s} (T_L s + 1)} \quad *$$

$$\frac{H_{on}}{H_o} = \frac{e^{-.1s}}{K_h e^{-.1s} (T_L s + 1) \cdot \frac{1}{(1+sT_N)(1+sT_I)}}$$

or

$$\frac{H_{on}}{H_o} = \frac{(T_N s + 1)(T_I s + 1)}{K_h (T_L s + 1)} \quad (3-9)$$

It is noted that the human-operator gain K_h can be lumped with the fixed elements L_n of (3-5). This results in a design for a single overall open-loop gain, K_T , from which the human-operator gain needed, K_h , or the plant gain needed, K_p , is determined upon specification of one or the other, i. e., $K_T = K_h K_p$. In a human-operator study the plant gain would be fixed at some nominal value since the investigation is focused upon the human operator aspects of the analysis. In general, this approach is used throughout, thus from (3-4)

$$1 + G'_p H_{on} \left[\frac{H_o}{H_{on}} \right] = 1 + G'_p H_{on} \left[\frac{K_h e^{-\tau s} (1 + s T_L)}{(1 + s T_N)(1 + s T_I)} \cdot \frac{1}{H_{on}} \right] = 0$$

$$= 1 + K_h K_p G'_p H_{on} \left[\frac{(1 + s T_L)}{(1 + s T_N)(1 + s T_I)} \right] = 0 \quad (3-10)$$

* The study regards only fixed values of τ and covers the entire range of τ variation by repetition over a sequence of values uniformly distributed, or otherwise, across $0.1 < \tau < 0.4$. Since the plots are in the Nyquist plane the additional burden is negligible and consists only of a phase shift in L_n , the magnitude of $e^{-j\omega\tau}$ being unity.

where the plant gain, K_p , has been removed from G'_p and H_{on} has been set equal to $e^{-\tau s}$. Equation (3-10), put into the form (3-5), is written

$$\frac{H_{on}}{H_o} + L_n = \frac{(1+sT_N)(1+sT_I)}{(1+sT_L)} + K_T G'_p e^{-\tau s} = 0 \quad (3-11)$$

Throughout the remainder of this work the expression for $\frac{H_{on}}{H_o}$ is written as

$$\Omega = \frac{H_{on}}{H_o} = \frac{(1+sT_N)(1+sT_I)}{(1+sT_L)} \quad (3-12)$$

and the human-operator gain is lumped into L_n as in (3-11). It should be noted that the L_n term of (3-11) has a FIXED PARAMETER structure, Fig. 3-1; each analysis considers only fixed values of τ across the range of interest, since the plots are in the complex plane the effect upon L_n consists only of an additional phase shift.

3.4. Stability Criterion

The standard Nyquist criterion for fixed parameter systems gives information about the characteristic equation of the closed-loop system. Thus, if the system function is

$$\frac{\theta_o}{\theta_i} = \frac{KP(s)}{1+KFP(s)}$$

then system stability is determined by whether the characteristic equation

$$1+KFP(s) = 0$$

has roots in the right-half plane. If the Nyquist plot of the open-loop transfer function encloses the -1 point, the number of such encirclements and their sense gives the number of zeros minus poles that $1+KFP(s)$ has in the right-half plane. If the system is open-loop stable, i. e., no poles of $KFP(s)$ in the right-half plane, the number of encirclements gives directly the number of zeros in the right-half plane. Thus, in this instance, any clockwise encirclement is sufficient to guarantee instability. If the system is open-loop unstable there are poles in the

right-half plane, and these must be accounted for in interpreting whether or not zeros are in the right-half plane. Clockwise encirclements are regarded as positive and counterclockwise encirclements as negative. For example, if

$$KFP(s) = \frac{K}{s(-1 + s\tau_1)}$$

the system is open-loop unstable (there is a right-half-plane pole). However, the Nyquist plot yields one clockwise encirclement, i.e.,

$$\text{number of zeros} - \text{number of poles} = Z - P = +1$$

but P is known to be unity thereby yielding $Z=2$. In this case, the system is closed-loop unstable, the characteristic equation has two roots in the right-half plane.

Can the Nyquist criterion be applied to systems containing time lags? Such a condition is always present with a human operator in the loop! The Nyquist criterion indicates the number of zeros of a function in the right half of the s plane by the number of revolutions the mapping of the function makes around a critical point of another complex plane for all values of s bounding the original right-half plane. The primary condition ensuring the validity of the Nyquist criterion is the analyticity of the function; since the exponential term representing the time lag effect is analytic in the region of interest the criterion is directly applicable without modification.*

The stability criterion used throughout this study follows from (3-11)

$$\Omega + L_n = \frac{(1 + s T_N)(1 + s T_I)}{(1 + s T_L)} + K_T G'_p e^{-\tau s} = 0 \quad (3-11)$$

* More detailed justification can be found elsewhere. (1)(2)(3)

- (1) Satche, M., Discussion, J. Appl. Mech., p. 419, Dec., 1949.
- (2) Dzung, L. H., The Stability Criterion, "Automatic and Manual Control", Tustin, A., Cranfield Conference papers, Butterworths (publishers), 1952.
- (3) Choksy, N. H., Time Lag Systems, "Progress in Control Engineering", Vol. 1, Academic Press, N. Y., 1962.

The stability of the system depends upon the number and type of encirclements the plot of (3-11) makes of the origin of its complex plane. * Variation in T_L , T_I and T_N causes Ω to take on the form of an area, however, and rotation of the "vector" representing (3-11) as the values of s traverse the bounds of the right-half plane must be interpreted with care. All points within the region of variation which result in a rotation of this "vector" such that instability is established represent combinations of values for the triple (T_L, T_I, T_N) which yield an unstable system. If no points exist which provide a stable result this reflects a situation where the human operator cannot stabilize (or control) the given system. The concept will be examined in detail later by building from simple examples. Thus, the region (3-12) is plotted for all possible values of the triple (T_L, T_I, T_N) ; the boundaries of the regions $\Omega = \frac{H_o}{H_o^n}(j\omega)$ (for various fixed ω each is called the "hull") are universally applicable for the selected range of variation on the time constants T_L, T_I, T_N . This range was obtained from empirical results published in the literature and was established by selecting the largest and smallest values that T_L, T_I and T_N ever assumed, as indicated from the many different experiments conducted to determine them. The universal plots may be used in conjunction with any system of interest. The first generalized version of this technique seems to have been given by Jones⁽¹⁾; regions of variation were not considered.

Rather than plot the sum (3-11) and observe any rotations about the origin, the following technique is used. The negative of L_n is plotted, i.e., $-L_n$, as well as the region of variation Ω ; Fig. 3-2(a) displays a hypothetical situation for explanatory purposes. If a vector is drawn from the $-L_n(j\omega)$ locus to the Ω

* The origin here plays the role of the "critical point".

(1) Jones, P., Stability of Feedback Systems Using Dual Nyquist Diagram, Trans. IRE Prof. Group on Circuit Theory, Mar., 1954.

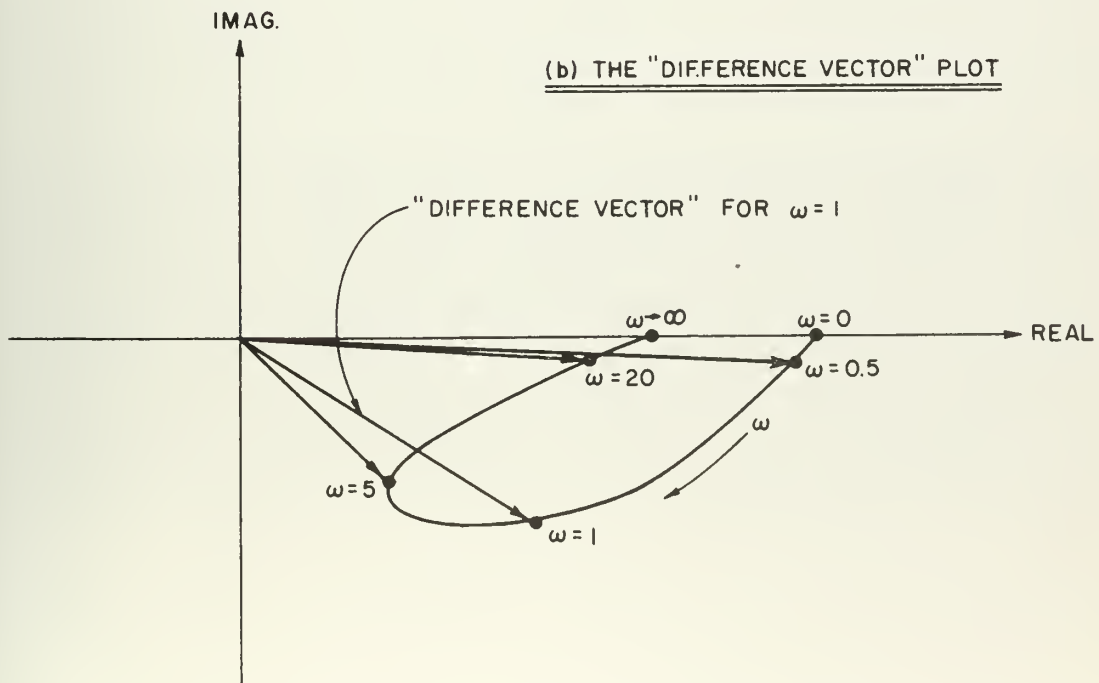
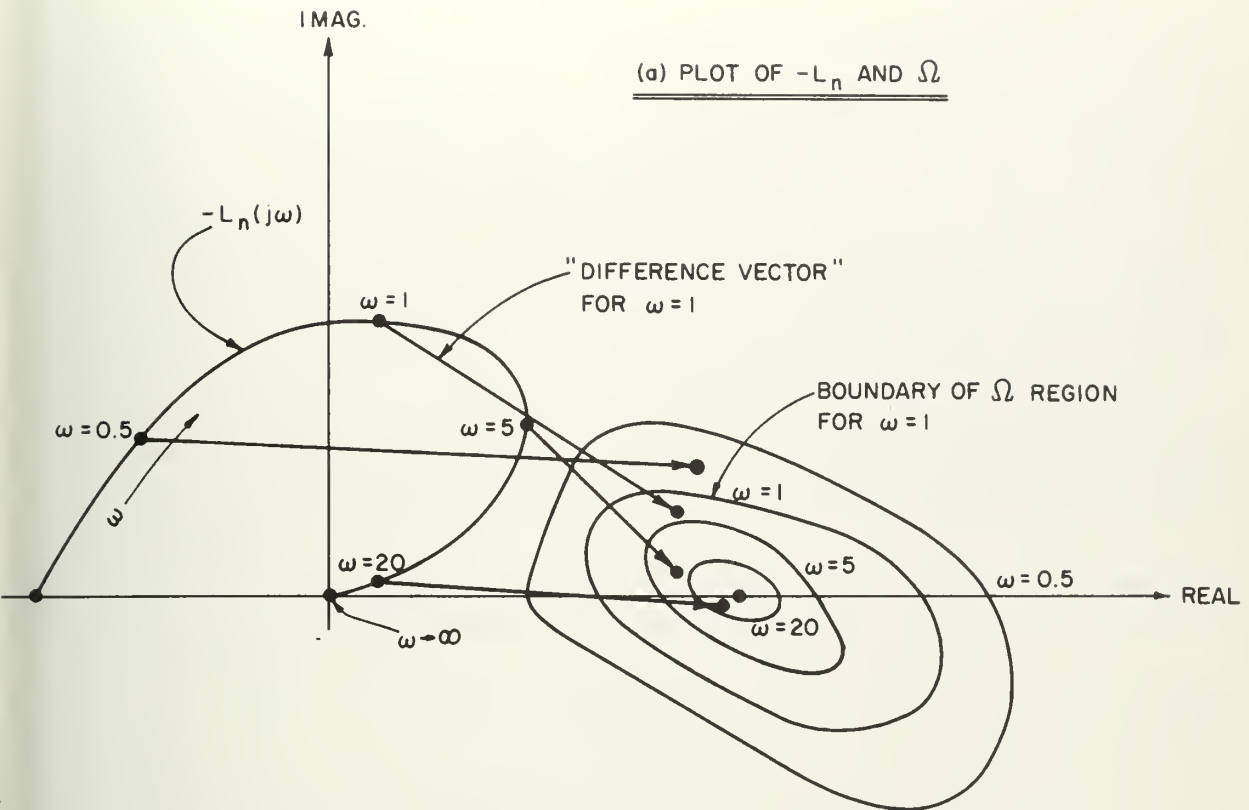


Fig. 3-2

"locus", each end of the vector being at the same value of ω , that "difference vector" is actually the function $\Omega + L_n$ and the rotations it experiences give stability information for that function. That is, stability is determined from the nature of the origin encirclements, if any, that occur in the "difference vector" plot, Fig. 3-2(b). The vectors shown in Fig. 3-2(a) are for specific values of the varying parameters as well as for particular frequencies. Thus, the locus of (3-11) may be visualized without carrying out the summation. The replacement of Ω by unity results in an ordinary Nyquist plot with the head end of the vector fixed at $+1+j0$ and the other tracing out the $-L_n(j\omega)$ locus. Thus, as a dividend the plot of $-L_n$ allows immediate observation of the plant-stability characteristics in those cases when $G_c=0$; the critical point from which stability is determined in such instances is at $+1+j0$. Examples will be offered in the next chapter.

It is not necessary, of course, to introduce the concept of the nominal human-operator transfer function H_{on} , since (3-3)

$$1 + H_o G'_p = 0 \quad (3-3)$$

could have been plotted in the form of either

$$H_o = -\frac{1}{G'_p} \quad (3-13)$$

or

$$-G'_p = \frac{1}{H_o} \quad (3-14)$$

and the "difference vector" between the two loci interpreted similarly to the explanation above. In (3-13) the design engineer using the techniques of this study would be required to plot the inverse frequency locus and in either (3-13) or (3-14) the general contours to be developed which define the region of variation would be complicated in their calculation by a linear phase shift due to the delay time. The use of the nominal expression for the human operator removes the delay from the varying portion of the plot and places the burden for its consideration on the FIXED PARAMETER portion, i. e., the user of the study always plots this locus since it

represents his particular system. His plot, however, need be modified by the addition of a linear phase shift to a single contour only instead of to a bounded region.

3.5 Summary

Under the assumption that the human operator can be described by a transfer function composed of five variable parameters, the system mathematics are manipulated so that three of these parameters (time constants) are displayed together. The remaining two parameters, gain and pure delay time, are lumped with the rest of the system; their effects upon the closed-loop stability characteristics are easily determined, since their influence upon the Nyquist contour of the plant is readily observed. It is shown that stability analysis of certain man-monitored system configurations also is possible. Stability estimates are reasonably accurate as long as the human-operator description used continues to represent the human controller adequately; this condition seems to hold for angular frequencies below 18 to 20 radians per second.

CHAPTER FOUR

THE STABILITY CRITERION AND THE VARIABLE PARAMETER MAPPING4.1 Introduction

One intent of this study is to provide a technique for viewing human controller-system response characteristics which is more intuitively satisfying than previous approaches. As a result of some interesting system transfer function manipulations, certain man-machine systems can be studied for stability on the familiar Nyquist plane. The time delay inherent in the human operator is easily handled and the effects of many of the human operator's physiological characteristics are plainly observed. The goal is to introduce a new approach to handling the quasi-linear human-operator transfer function without excluding the enriching properties of human adaptive range and pure time delay.

In this chapter a complete analysis of a very simple single-varying-parameter system is conducted to provide familiarity with how the theory discussed in Chapter Three actually is used. Following this presentation, the multidimensional variable parameter mapping is discussed in detail. The mapping function is that portion of the human-operator transfer function which previously was isolated as a separate term (the three variable time constants). Certain properties of the mapping which relate to necessary and sufficient conditions on system stability are considered. Finally, plots of the mapping are presented (also on transparent overlays); it is these plots which are used in Chapter Five to predict human control capabilities.

4.2 Analysis of a Simple System

To facilitate an understanding of the modified Nyquist diagram used in this study consider the system depicted in Fig. 4-1. The root loci shown in Figs. 4-2 and 4-3 are representative of this system, i. e., for the same open-loop gain $K = K_1$ it is possible to have values for the variable time constant x such that a stable closed loop occurs, Fig. 4-2; or an unstable closed loop prevails, Fig. 4-3. The encircled crosses depict closed-loop poles while the others represent open-loop

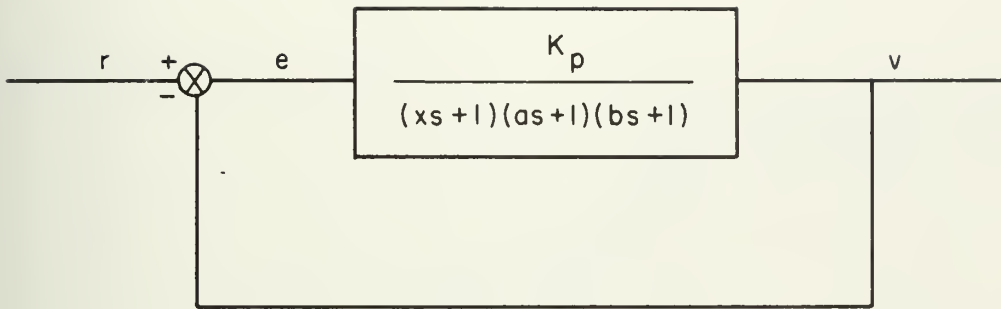


Fig. 4-1 : A System With One Variable Parameter

poles. Such a system, with only one varying parameter, can be studied by way of example to gain an understanding of the dual Nyquist plotting technique. Generalization to a three varying parameter mapping follows easily.

A specific example of Fig. 4-1 where $b=1$ and $a=\frac{1}{2}$ follows. The closed-loop function has a denominator of the form

$$D(s) = 1 + \frac{K_p}{(sx+1)(\frac{s}{2}+1)(s+1)} = 1 + \frac{2K_p/x}{(s+\frac{1}{x})(s+1)(s+2)}$$

If $x = \frac{1}{4}$ then

$$F(s) = \frac{8K_v}{(s+4)(s+1)(s+2)} = \frac{K}{(s+4)(s+1)(s+2)}, \quad K = 8K_v$$

Therefore,

$$D = 1 + F = s^3 + 7s^2 + 14s + 8 + K = 0$$

is the characteristic equation. From the Routh array, the stability limit for the root locus gain is found

| | | |
|-------|-----------------------|---------|
| s^3 | 1 | 14 |
| s^2 | 7 | $(8+K)$ |
| s^1 | $\frac{-(8+K)+98}{7}$ | |
| s^0 | | |

to be $K=90$, $\therefore K_v = \frac{90}{8} = 11.25$; the point at which the closed-loop poles cross the $j\omega$ axis is

$$7s^2 = -98$$

$$s^2 = -14$$

$$\therefore s = \pm j3.74$$

This situation is shown in Fig. 4-4 by the solid line loci. Now suppose, $x = \frac{1}{3}$

instead of $\frac{1}{4}$ then

$$D(s) = 1 + \frac{6K_v}{(s+3)(s+1)(s+2)} = 1 + \frac{K}{(s+3)(s+1)(s+2)}$$

and

$$D = 1 + F = s^3 + 6s^2 + 11s + 6 + K = 0$$

is the characteristic equation. Again, using the Routh array, the stability limit is found

$$\begin{array}{c|cc} s^3 & 1 & 11 \\ s^2 & 6 & (6+K) \\ s^1 & \frac{-(6+K)+66}{6} & \\ s^0 & & \end{array}$$

to be $K=60$, $\therefore K_v = \frac{60}{6} = 10$; the crossing of the $j\omega$ axis by the closed-loop poles occurs at

$$6s^2 = -66$$

$$s^2 = -11$$

$$\therefore s = \pm j3.32$$

The loci for this system configuration are shown dotted in Fig. 4-4 as is the pole at $-\frac{1}{x} = -3$.

If x is allowed to vary over the arbitrarily selected range $0 \leq x \leq 5$ the system will be characterized by both stable and unstable response properties depending upon the value x assumes, provided the gain is "properly" fixed. Thus, for $K=90$ the system is at its stability limit for $\frac{1}{x} = 4$, and is surely unstable when $\frac{1}{x} = 3$ for the same gain, since the gain driving the closed-loop poles onto the imaginary axis for $\frac{1}{x} = 3$ is only $K=60$. Thus, at $K=90$, there is system

stability for $\frac{1}{x} > 4$ and instability for $\frac{1}{x} \leq 4$, see Fig. 4-5.

The plot of these situations using the dual Nyquist approach is now considered. Again, letting

$$D(s) = 1 + \frac{11.25}{(sx+1)(s+1)(\frac{s}{2}+1)}$$

and with the previous knowledge that

$$K_v = 11.25 \quad \text{and} \quad x = \frac{1}{4}$$

corresponds to the stability limit, while

$$K_v = 11.25 \quad \text{and} \quad x = \frac{1}{3}$$

corresponds to instability, the denominator of the closed loop function, $D(s)$ is rewritten.

$$D(s) = 1 + V(s)R(s)$$

The $V(s)$ portion will contain the varying parameters and the $R(s)$ portion the fixed ("rigid") parameters.

Thus,

$$V = \frac{1}{sx+1} \tag{4-1}$$

and

$$R = \frac{11.25}{(s+1)(\frac{s}{2}+1)} \tag{4-2}$$

so

$$D = 1 + RV \left[\frac{V_n}{V_n} \right] \tag{4-3}$$

The nominal function, V_n , is taken as unity, i. e., $x=0$, in the expression for V . Then the roots of

$$D = 1 + RV \left[\frac{V_n}{V_n} \right] = 0$$

are those of

$$\frac{V}{V_n} + RV_n = 0 \tag{4-4}$$

or

$$(sx+1) + \frac{11.25}{(s+1)(\frac{s}{2}+1)} = 0 \quad (4-5)$$

In equation (4-5), the expression (4-2) plays the role of L_n in (3-5). Thus,

$$L_n = \frac{11.25}{(s+1)(\frac{s}{2}+1)} \quad (4-6)$$

and similarly,

$$\frac{V_n}{V} = (sx+1) \quad (4-7)$$

plays the role of $\frac{H_{on}}{H_o} = \Omega$.

In analogy to the method described in Chapter Three, the expressions for, $-L_n$, representing the fixed parameter contour and, $\frac{V_n}{V}$, the region of criticality (here a straight line) representing the varying parameter structure, are plotted. It is noted that this varying parameter structure is characterized by the mapping $\frac{V_n}{V}$ over the range of x and for specific frequencies ω . The remainder of the example clarifies this notion.

The expression (4-7) plots as a straight line for any given ω with its bounds defined by the end points on the range of x . For example,

$$\text{let} \quad s = j\omega = j \quad \text{i. e.} \quad \omega = 1$$

$$\text{then} \quad \frac{V_n}{V}(j1) = jx + 1$$

$$\text{At } x=0 \quad \frac{V_n}{V}(j1) = 1$$

$$\text{At } x=5 \quad \frac{V_n}{V}(j1) = j5 + 1$$

For $\omega = 0.5$:

then

$$\frac{V_n}{V}(j0.5) = j(0.5)x + 1$$

At $x = 0$

$$\frac{V_n}{V}(j0.5) = 1$$

At $x = 5$

$$\frac{V_n}{V}(j0.5) = j2.5 + 1$$

For $\omega = 2$:

then

$$\frac{V_n}{V}(j2) = j2x + 1$$

At $x = 0$

$$\frac{V_n}{V}(j2) = 1$$

At $x = 5$

$$\frac{V_n}{V}(j2) = j10 + 1$$

The results are plotted in Fig. 4-6 and are represented by the vertical line running from $1+j0$ to $1+j\infty$ as $\omega \rightarrow \infty$. Actually shown on Fig. 4-6 are the situations reflecting $\omega = 0.5$, $\omega = 1$ and $\omega = 2$. Thus, when $\omega = 1$, the $\frac{V_n}{V}$ plot may assume any point along the vertical line segment from $1+j0$ to $1+j5$ depending upon the particular value of x . This segment is defined in Fig. 4-6 by the vertical distance dimensioned $\omega = 1$, and the others similarly.

A plot for the FIXED PARAMETER structure, $-L_n$, is obtained from the negative of (4-6). It is easy to see that

$$-L_n(j0) = -11.25 = 11.25 \angle 180 = 11.25 \angle -180$$

and that the

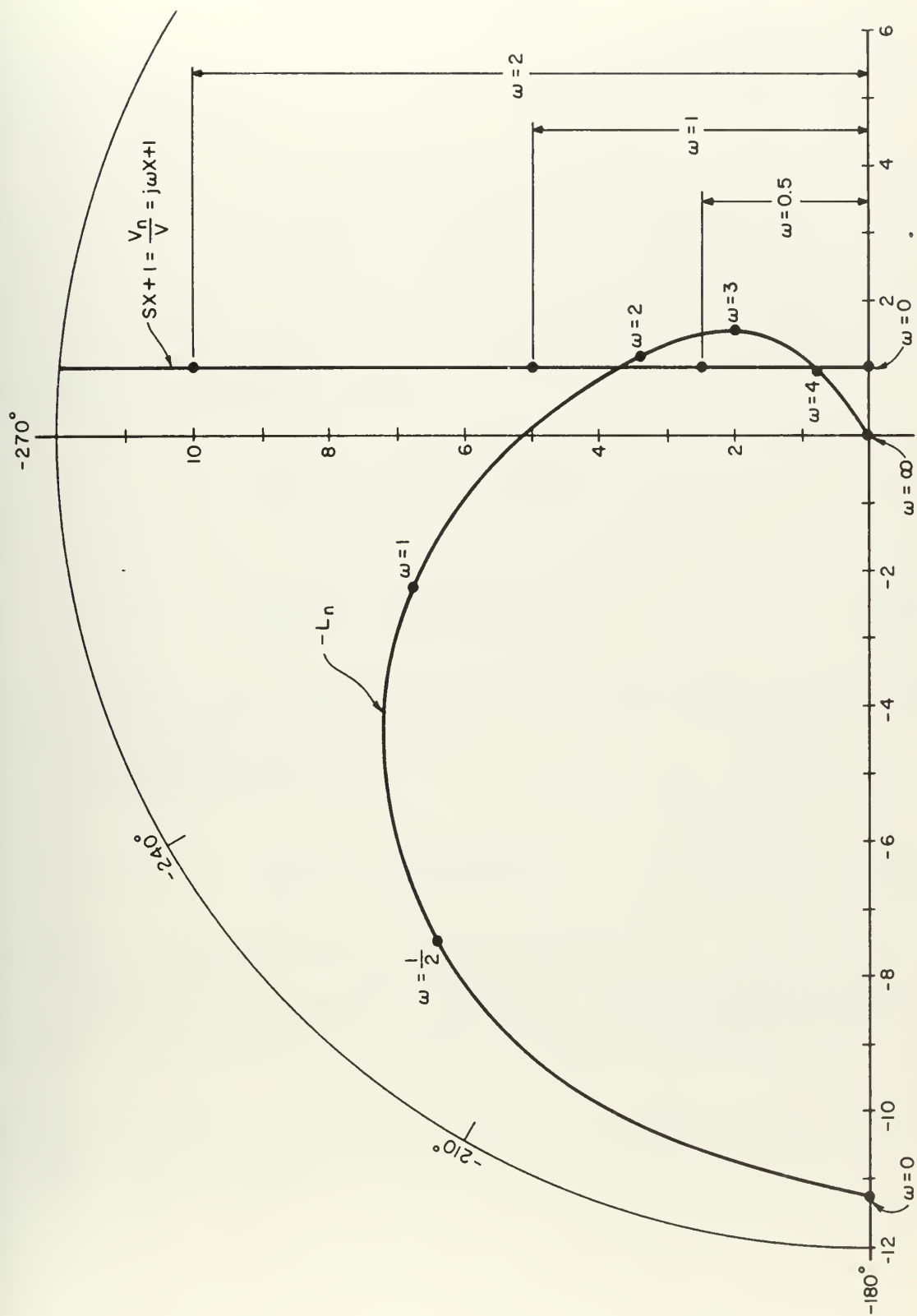


Fig. 4-6 : Dual Nyquist Plot of a Varying-Parameter System

$$\lim_{\omega \rightarrow \infty} [-L_n(j\omega)] = 0 \angle 0$$

thus providing the end points of the contour. Other points may be found in the normal, tedious fashion or by transference from a Bode' plot. This contour is also shown in Fig. 4-6.

It is difficult to see at this point that the "difference vector" drawn from, $-L_n$, to the $\frac{V_n}{V}$ contour (head end at the latter), and connecting identical values of ω on each contour, does itself undergo a clockwise rotation if $x \geq \frac{1}{4}$. Without some experience viewing the plot, initial interpretation may be formidable; for this reason, consider x assuming only two specific values in its range. When $x = \frac{1}{5}$, the system is stable, and when $x = \frac{1}{3}$, the system is unstable; these two conditions are now studied. Thus, for

$$F_{\frac{1}{5}}(s) = \frac{11.25}{(\frac{s}{5} + 1)(s + 1)(\frac{s}{2} + 1)} \quad (4-8)$$

and

$$F_{\frac{1}{3}}(s) = \frac{11.25}{(\frac{s}{3} + 1)(s + 1)(\frac{s}{2} + 1)} \quad (4-9)$$

there is stability with (4-8) and instability with (4-9). This can be checked quickly by counting sign changes in a Routh test.⁽¹⁾

It should be recalled that

$$F_{\frac{1}{4}}(s) = \frac{11.25}{(\frac{s}{4} + 1)(s + 1)(\frac{s}{2} + 1)} \quad (4-10)$$

represents a system precisely at its stability limit.

Separation of the varying (in this instance, "fixed", but arbitrary) and fixed parts of the characteristic equation results in (4-5), repeated below.

⁽¹⁾ Gardner, M. F. and J. L. Barnes, "Transients in Linear Systems", Vol. I, Wiley, New York, 1942.

$$(sx + 1) + \frac{11.25}{(s + 1)(\frac{s}{2} + 1)} = 0 \quad (4-5)$$

where $x = \frac{1}{3}$ or $\frac{1}{5}$.

Thus, for $x = \frac{1}{5}$:

$$\overbrace{\quad L_n \quad} + \frac{11.25}{(s + 1)(\frac{s}{2} + 1)} = 0$$

The $-L_n$ plot is the same as in the Fig. 4-6 plot and is reproduced in Fig. 4-7; the $(\frac{s}{5} + 1)$ plot is again a straight line (this time for varying ω and fixed x) upon which are marked, to the right of the line, the corresponding frequency points. It is noted in Fig. 4-7, that the difference vector (shown for $\omega = 1$) never undergoes a complete encirclement (360° or more). At $\omega = 0$, the difference vector points toward 0° (due right), while at slightly less than $\omega = 2$ (first intersection of $-L_n$ and $\frac{V_n}{V}$), the difference vector points toward -90° (straight down). Somewhere in the neighborhood of $\omega = 3$, this vector experiences its largest clockwise excursion (approximately -120°), after which it begins a counterclockwise return rotation until at $\omega = 4$ it has almost returned to 0° . As $\omega \rightarrow \infty$, it undergoes a counter-clockwise rotation approaching 90° in the limit. Clearly, a complete encirclement does not occur.

For $x = \frac{1}{3}$:

The $-L_n$ plot is, of course, unchanged. The $(\frac{s}{3} + 1)$ plot superposes over the line describing $(\frac{s}{5} + 1)$ but with different frequency correspondences. These are shown by the markings which appear to the left of the vertical line in Fig. 4-7. It can be seen from Fig. 4-7 that as ω runs from $\omega = 3$ to $\omega = 4$ the difference vector between the two contours undergoes an encirclement. It is clear that whether or not encirclement occurs depends upon the relative values of frequency on the $-L_n$ and $(\frac{s}{3} + 1)$ contours.

A magnification of the situation is contained in the enlarged sketches of Fig. 4-8. In Fig. 4-8(a) is shown the hypothetical situation where ω_{-L_n} (at intersection with $\frac{V_n}{V}$) = 3.7 and $\omega_{\frac{V_n}{V}}$ (at intersection with $-L_n$) = 4.7. The effects of

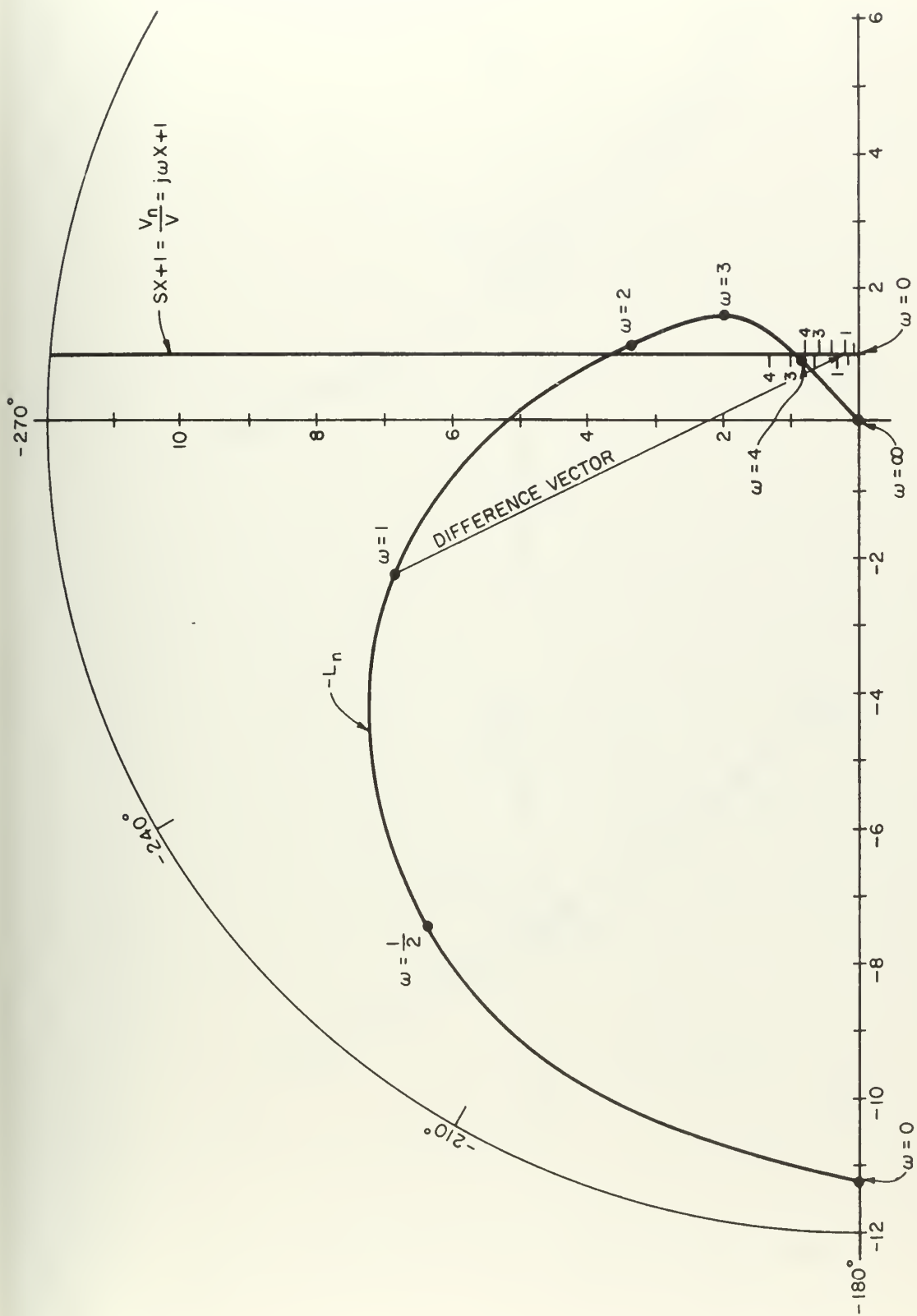


Fig. 4-7 : Detailed Dual Nyquist Plot

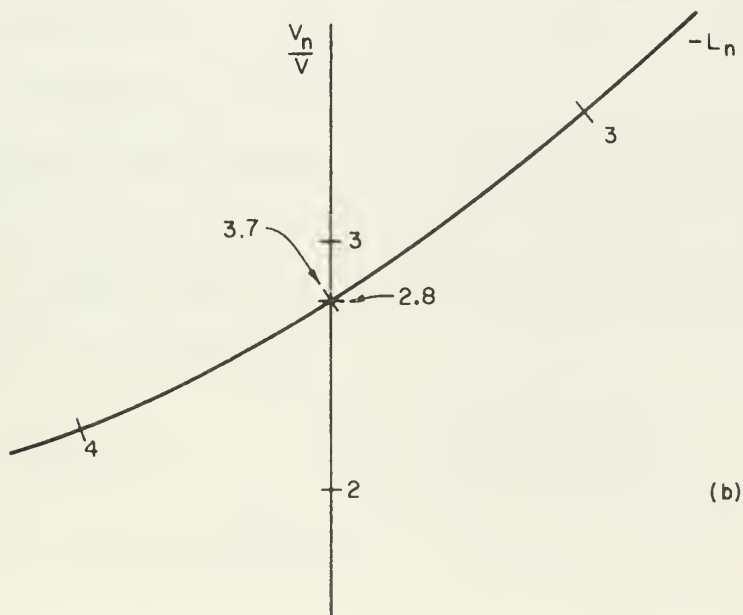
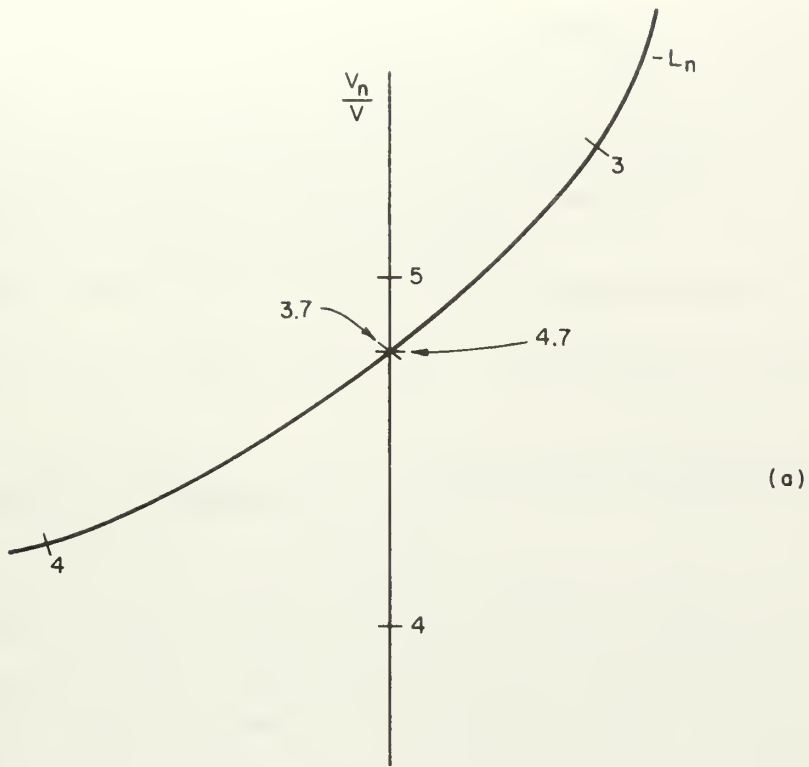


Fig. 4-8 : Hypothetical Crossing Situations (in the small)

encircling the right half of the s -plane upon the mapping of the difference vector between $-L_n$ and $(sx+1)$ is shown in Fig. 4-9(a) and (b). In a qualitative plot of the difference-vector's angular rotation, Fig. 4-10, there is no encirclement of the origin and the system therefore has no right-half-plane roots in its characteristic equation; since the open-loop system is stable this conclusion is permissible. Thus, the system is closed-loop stable.

A different state of affairs is depicted in Fig. 4-11(b) where ω_{-L_n} (intersection) = 3.7 and $\omega_{\frac{V_n}{V}}$ (intersection) = 2.8. The effect upon the mapping of the

difference vector is such that, Fig. 4-11(a), (b) and the qualitative plot of Fig. 4-12, two encirclements occur indicating a system with two closed-loop poles in the right-half plane and unstable characteristics.

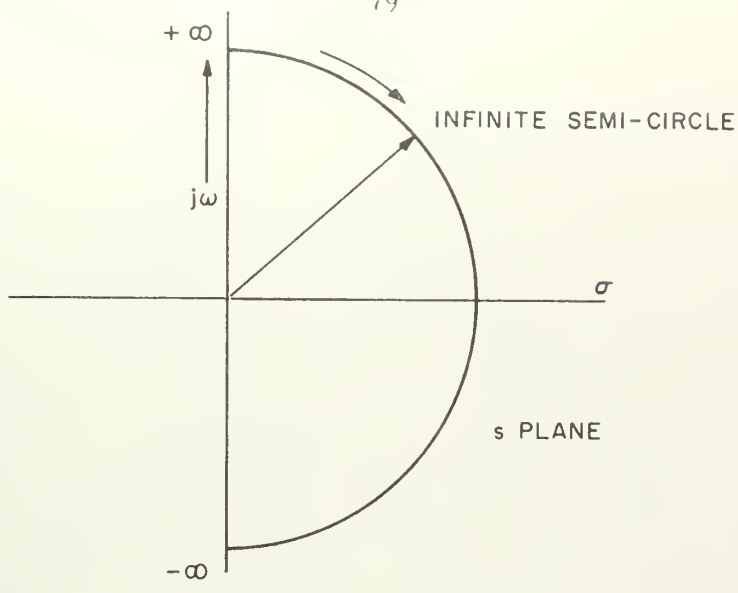
In this simple case, it can be inferred that if the value of the frequency on the $-L_n$ contour is greater than that on the locus of the varying structure, at the critical intersection point, then instability occurs. In short, if the $-L_n$ locus "encloses" the varying locus there is an unstable condition; "encloses" is interpreted to mean: if the frequency at the intersection of the two loci is ω_i on $-L_n$, and if in approaching the point $-L_n(j\omega_i)$ as the $-L_n$ contour is traversed in the direction of increasing frequency, the point $\frac{V_n}{V}(j\omega_i)$ is passed to the right of the $-L_n$ contour (called enclosure), then the system is unstable. Again, there is inherent here the assumption of open-loop stability.

Before this simple example is left, the precise values of the intrinsic parameter ω (at the points of intersection) are calculated for each contour. This is done for completeness of the example. It is not required for the stability test under ordinary circumstances since, with practice, encirclement can be envisioned from the relative locations of discrete frequency points on the dual Nyquist plot.

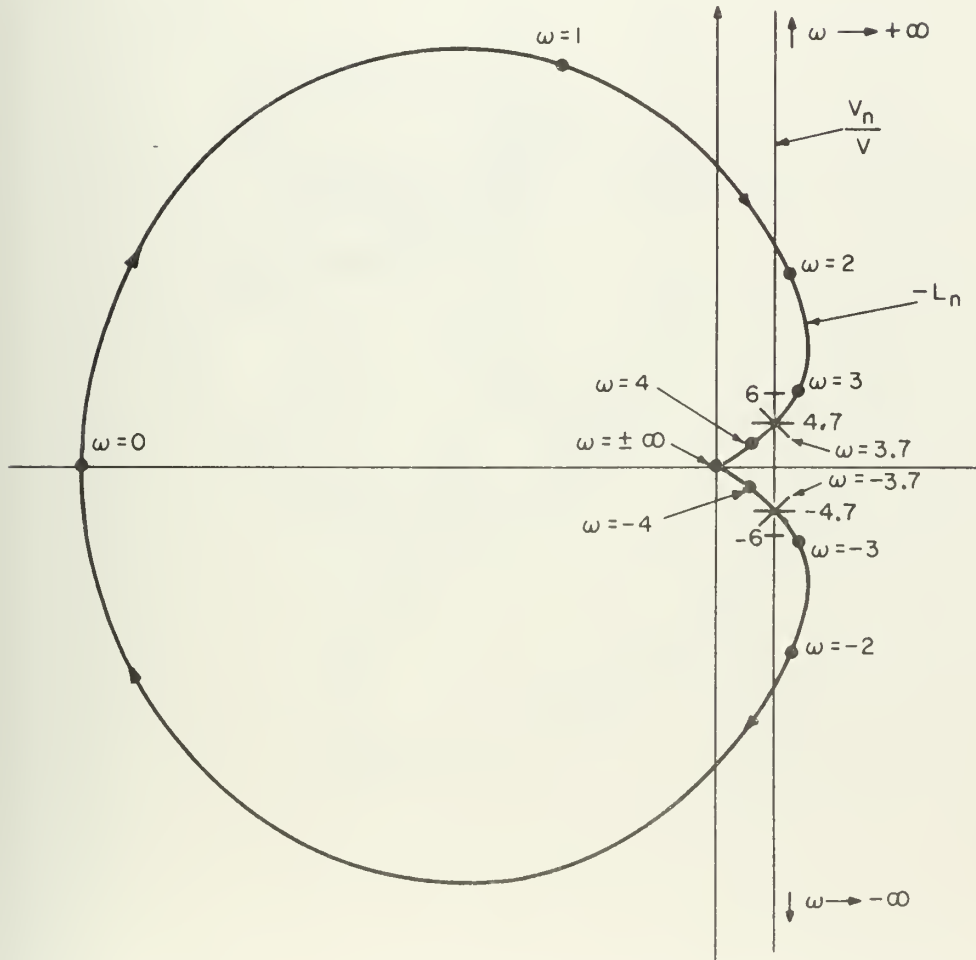
For the stable configuration:

$$\frac{V_n}{V} = \frac{s}{5} + 1 \qquad -L_n = \frac{-11.25}{(s+1)(\frac{s}{2} + 1)}$$

Intersection occurs in this example when

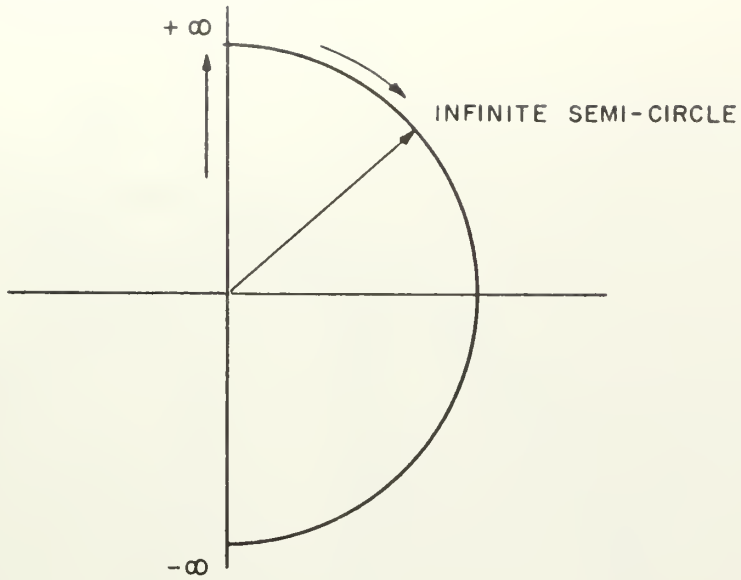


(a)



(b)

Fig.4-9 : Hypothetical Crossing Situations (in the large)



(a)

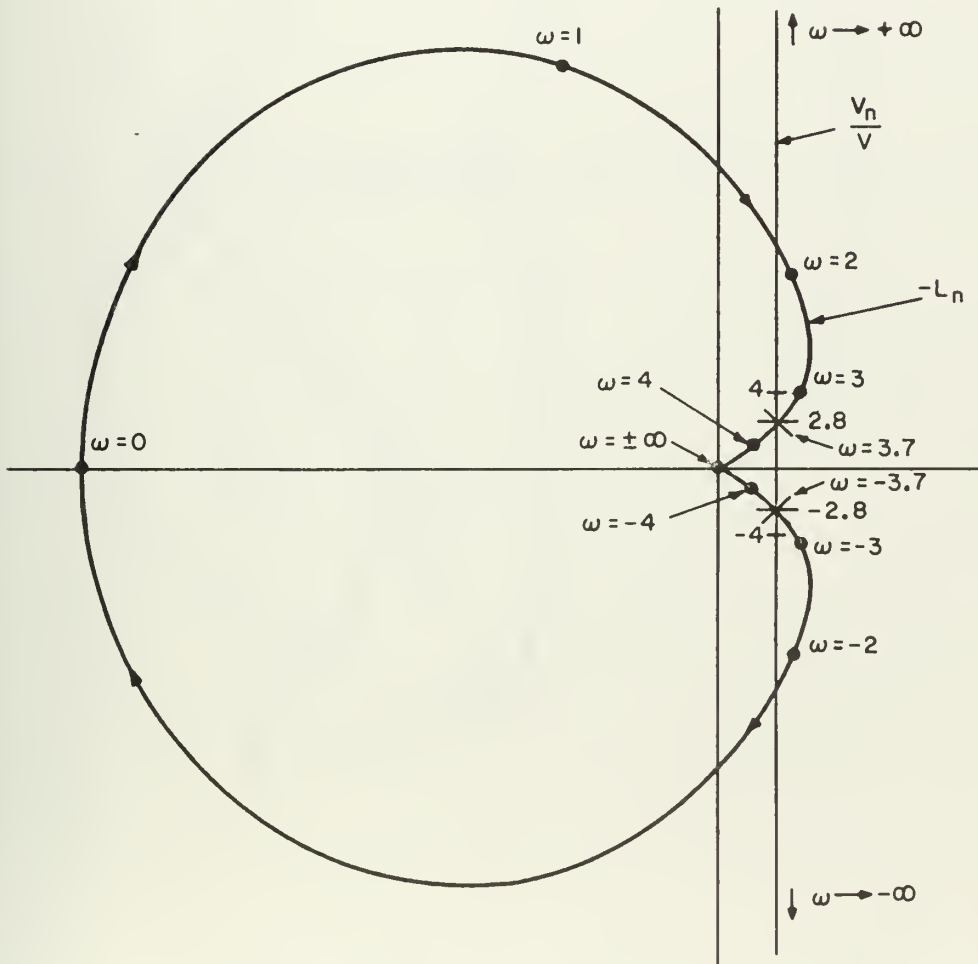


Fig. 4-11 : Hypothetical Crossing Situation (in the large)

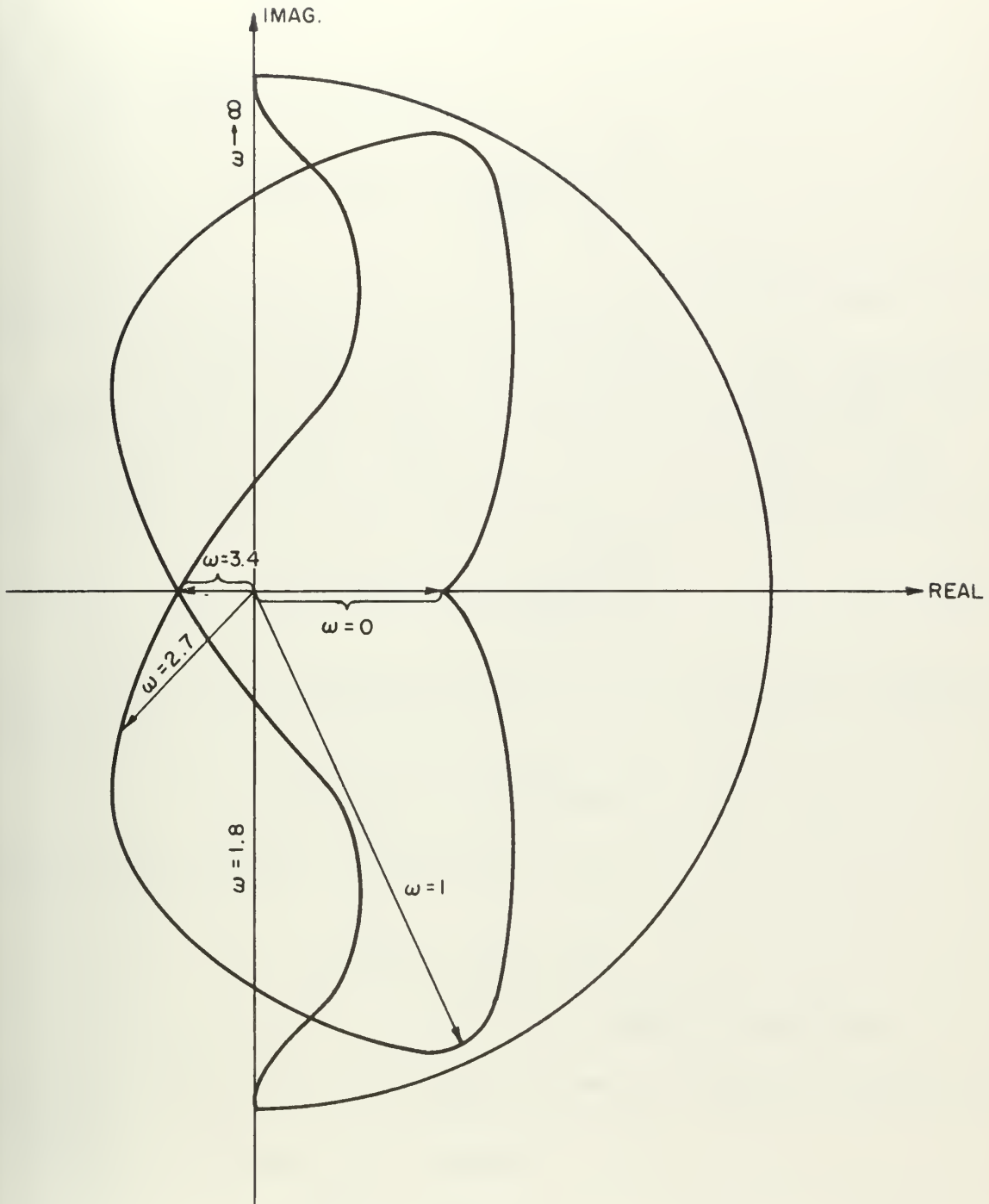


Fig. 4-12 : Excursion of the Difference Vector

$$\operatorname{Re} \frac{V_n}{V}(j\omega) = \operatorname{Re} [-L_n(j\omega)] = 1$$

where "Re" stands for "the real part of".

Thus, from

$$\operatorname{Re} [-L_n(j\omega)] = \frac{-11.25 \left(1 - \frac{\omega^2}{2}\right)}{\left(1 - \frac{\omega^2}{2}\right)^2 + \frac{9}{4}\omega^2} = 1$$

it is found that $\omega_1 = 1.87$ and $\omega_2 = 3.75$, where the subscripts indicate the frequencies of first and second intersection respectively.

It is now clear (since in this simple case the locus of the varying portion is the same straight line but with different frequency markings) that, when $x = \frac{1}{5}$, $-L_n(j3.75)$ does not "enclose" $\frac{V_n}{V}(j3.75)$; when $x = \frac{1}{3}$, $-L_n(j3.75)$ does "enclose" $\frac{V_n}{V}(3.75)$. Of course, this statement can be checked analytically. Thus,

$$\operatorname{Im} [-L_n(j\omega)] = \frac{-11.25 \left(-\frac{3}{2}\omega\right)}{\left(1 - \frac{\omega^2}{2}\right)^2 + \frac{9}{4}\omega^2} \quad (4-11)$$

where "Im" signifies "the imaginary part of". For $\omega = 3.75$, (4-11) evaluates to 0.936. But

$$\operatorname{Im} [-L_n(j\omega)] = \operatorname{Im} \left[\frac{V_n}{V}(j\omega) \right]$$

at the point of intersection. Therefore, for $x = \frac{1}{5}$

$$\operatorname{Im} \frac{V_n}{V}(j\omega) = \frac{\omega}{5} = 0.936$$

or

$$\omega = 4.68, \text{ on } \frac{V_n}{V}(j\omega) \text{ at intersection}$$

and, clearly, $\frac{V_n}{V}(j3.75)$ is to the left of the contour $-L_n(j\omega)$, indicating a stable system.

For the case when $x = \frac{1}{3}$, similarly,

$$\operatorname{Im} \frac{V_n}{V}(j\omega) = \frac{\omega}{3} = 0.936$$

or

$$\omega = 2.81, \text{ on } \frac{V_n}{V}(j\omega) \text{ at intersection}$$

and, clearly, $\frac{V_n}{V}(j3.75)$ falls to the right of the contour $-L_n(j\omega)$, indicating an unstable system.

A thumb rule for interpreting the stability criterion in this simple situation has now been established. In the original example, Fig. 4-7, an attempt at interpretation can be made using the additional insight just gained. It is known that the intersection of interest occurs at the point $-L_n(j3.75)$, and instability obtains for any value of x such that

$$\omega \frac{V_n}{V} (\text{intersection}) < \omega_{-L_n} (\text{intersection}).$$

In the particular case where the locus of the varying portion is a straight line there is the condition that

$$\text{Im}(sx+1) = x \frac{V_n}{V} = 0.936$$

But for stability,

$$x \frac{V_n}{V} > 3.75$$

therefore, since

$$x = \frac{.936}{\frac{V_n}{V}}$$

for stability

$$x < \frac{.936}{3.75} = .250$$

Thus, for stability, $x < .25$ which reflects precisely the conditions seen in the root locus plots of Fig. 4-4 and 4-5.

4.3 Multidimensional Variable Parameter Mapping

The system examined in Section 4.2 had but a single varying parameter.

It is for this reason that the locus of the critical point governing stability moved along a single curve, in that example, a straight line. When more than one parameter in a system varies (multidimensional parameter variation), the locus of critical points assumes the form of an area. In a well-behaved situation, for example, a mapping of the boundaries of varying α and β , Fig. 4-13(a), appears as in Fig. 4-13(b). In this case, every element XY which scans the constrained range of $\alpha\beta$ space is inside the mapping's periphery $0'1'2'3'$, called the "hull". In such a case the mapping is "complete" and with a suitable generalization of Section 4.2 the stability criterion is necessary and sufficient.

In the general case, however, end point mapping of the region of variation produces results as appear in Figs. 4-14(a) and (b). Here, for a given distance along the α -axis the element $X_a Y_a$ maps into a line contained inside the hull; but, further along it is clear that $X_b Y_b$ has points which are not contained in the hull, e.g., $X'_b Y'_b$, Fig. 4-14(b). Thus, although the end points of the $\frac{V_n}{V}$ line segments are the images of the end points of the rectangular region in $\alpha\beta$ space the mapping may not be "complete". It is clear, however, that all elements $X'Y'$ are inside the triangle $0'1'3'$; this "polygon" is called the "convex hull". Obviously, the convex hull contains points which are not a result of mapping the $\alpha\beta$ variation. It is for this reason that exclusive use of the convex hull provides only a sufficient condition for stability, i.e., for no "encirclement" a system is stable but, since an "encirclement" might be of a point or points not part of the mapping, an "encirclement" does not necessarily indicate "instability". This becomes clear in the examples. It should be noted, however, that a complete mapping can always be obtained. One need only plot a few points such as $X'_b Y'_b$ of Fig. 4-14(b) and the true hull can be sketched quickly. It is shown later that a closed form expression (a parabola) can be obtained for the envelope of the lines $X'Y'$ in Fig. 4-14(b); this will have direct application in the human operator situation.

Frequently, the requirement for mapping additional points is precluded by the relative location of the fixed and varying contours. For example, assuming the situation shown in Fig. 4-15 holds in a similar way for all w of pertinence in addition to $w=w_1$, it is obvious that whether $0'1'2'3'$ or $0'1'2'2''3'$ is used

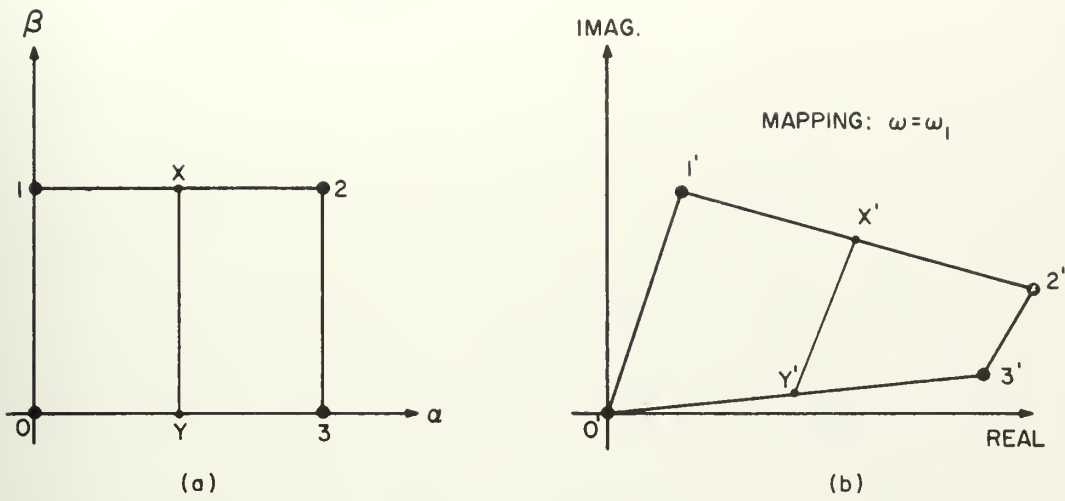
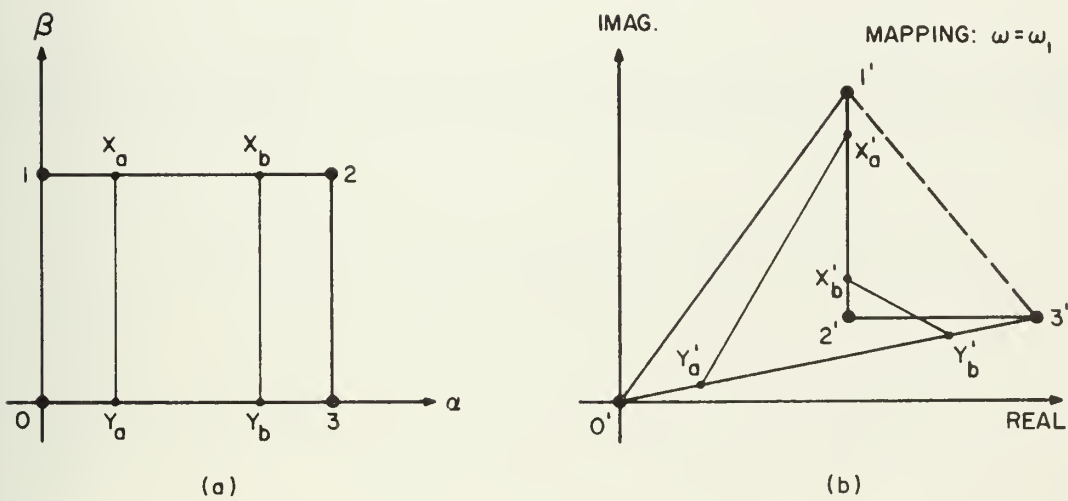


Fig. 4-13 : Boundary Mapping Resulting in Complete Hull

Fig. 4-14 : Boundary Mapping with Hull which is Not Complete
(also shows Convex Hull)

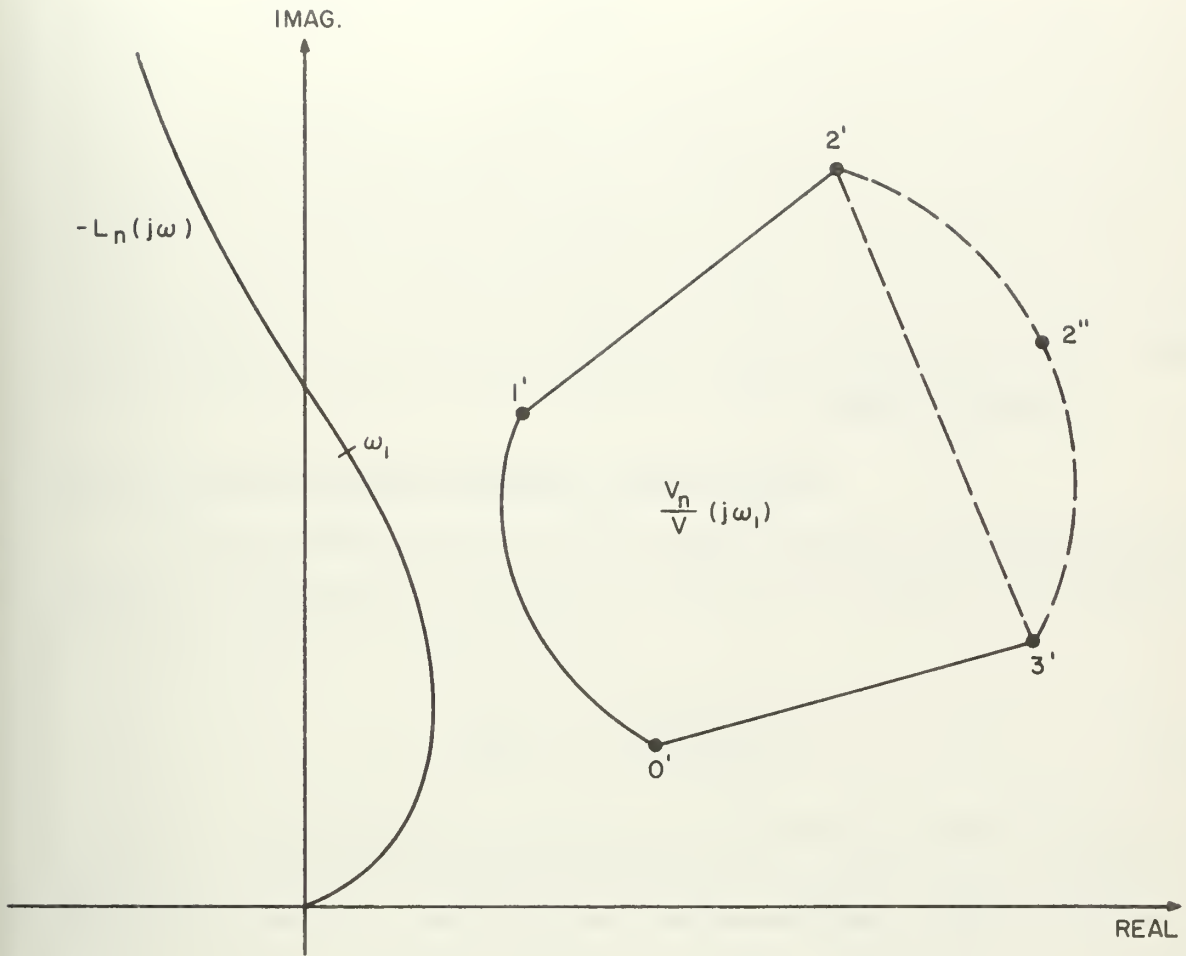


Fig. 4-15 : Hypothetical Variable Parameter Situation

makes little difference. The shaping of $-L_n$ is carried out with respect to a region common to both. This property does not hold dogmatically when the human operator is considered, since being an adaptive mechanism, he will strive to assume characteristics in that part of his "varying region" which afford stability. It is through this understanding, however, that insight may be gained concerning the amount of effort the human operator has to expend to maintain control. At times it is possible to demonstrate that a mapping of the end points in $\alpha\beta$ space yields almost the true hull and is negligibly different from the complete mapping, so that practical use can be made of the end point mapping. This is the situation with the human operator and the mapping is called "almost-complete".

4.4. A Map of the Human-Operator Adaptive Range

With the human-operator's gain-varying capability and delay time lumped into the fixed parameter structure, the varying portion of the characteristic equation is given by (3-12), repeated here

$$\Omega = \frac{H_{on}}{H_o} = \frac{(1 + sT_N)(1 + sT_I)}{(1 + sT_L)} \quad (3-12)$$

In order to study the human operator pure delay effect, τ , over its entire range of variation, several plots for different values of τ must be made. Usually, due to the limited range of τ and the prospect of interpolating middle range effects from the end bounds on τ , one or two plots will suffice. The plots (of $-L_n$, (3-11)) are all identical except for the phase lag introduced by the particular value of τ used.

The region of variation for T_L , T_I and T_N is shown in Fig. 4-16, i. e., equation (3-12) is used to map this three-dimensional volume into the complex plane plot of $\Omega(j\omega)$. The region covered by Ω as the triple (T_L, T_I, T_N) assumes all values on the bounds of the rectangular parallelipiped is shown in Fig. 4-17. (Later the completeness of this "end point mapping" is considered and it is shown that the end point hull is almost complete as described in Section 4.3.) For motion along the boundary AB of Fig. 4-16 there is obtained

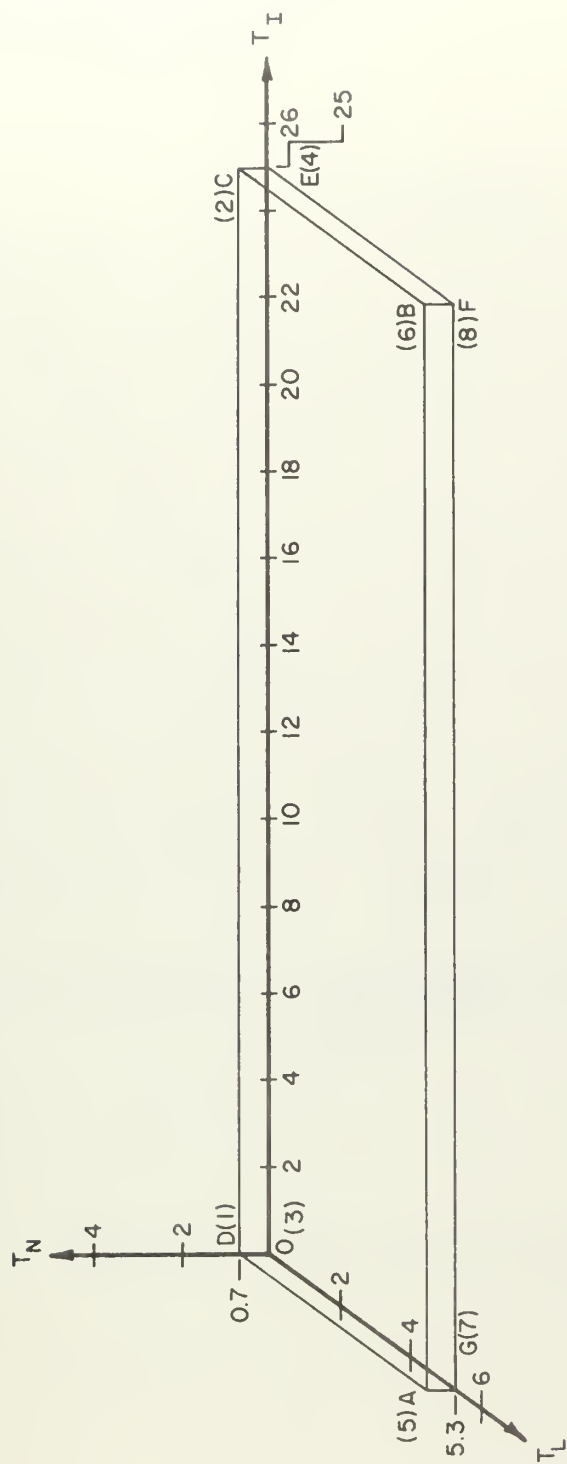


Fig. 4-16 : Human Operator Adaptive Range

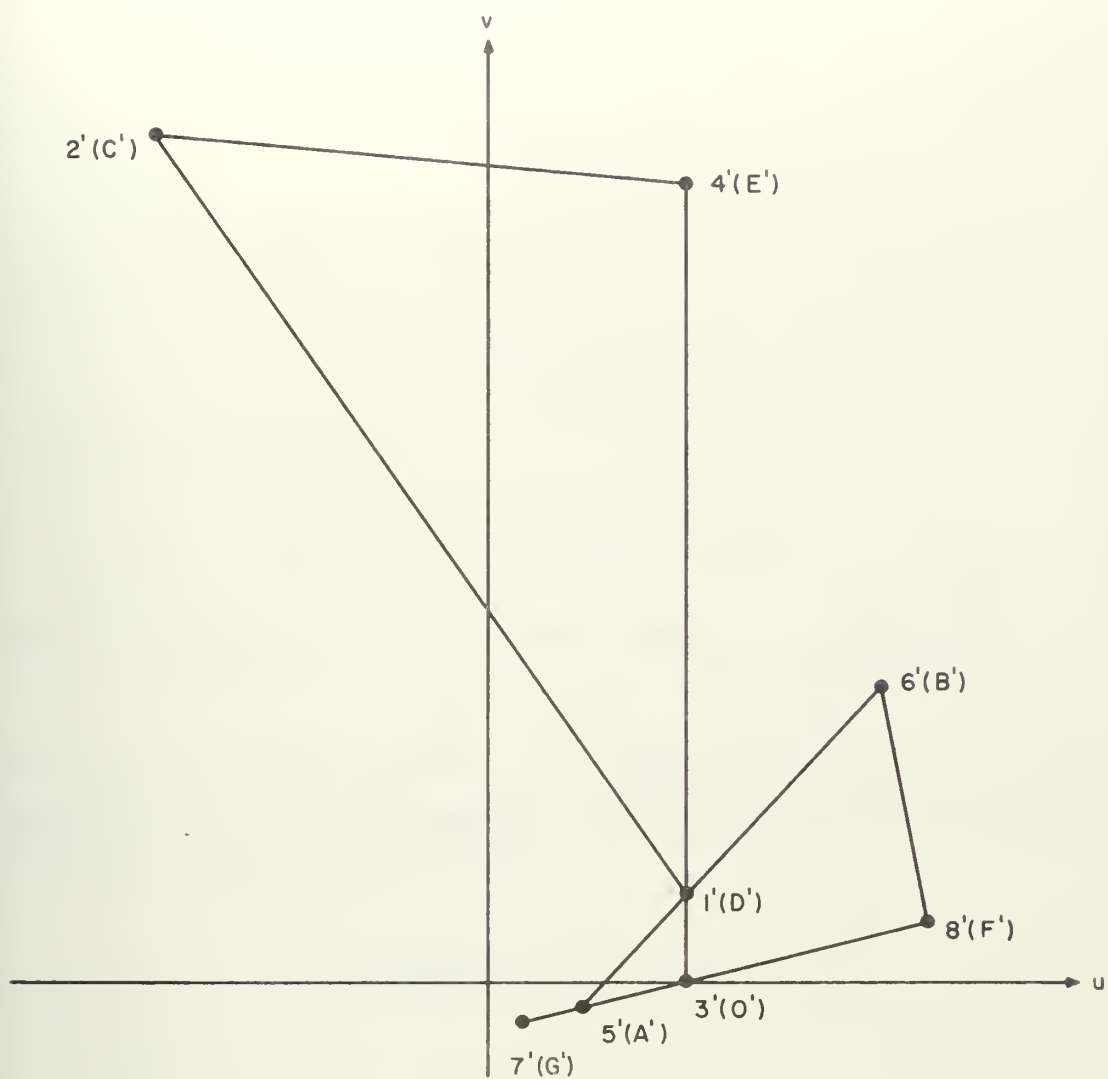


Fig. 4-17 : Mapping of Human Operator Adaptive Range

$$\Omega = \frac{(T_I s + 1)(0.7s + 1)}{(5.3s + 1)} \quad (4-12)$$

but for a fixed ω this is the equation of a straight line in T_I . For example, let $\omega = 1$, then

$$\Omega(j1) = \frac{(T_I j + 1)(0.7j + 1)}{(5.3j + 1)}$$

and after some manipulation

$$\Omega(j1) = \frac{T_I(4.6 + j4.71) + (4.71 - j4.6)}{29}$$

which is recognized as the general vector equation of a straight line

$$\bar{R} = \bar{A} + t\bar{B}$$

Since T_I has a range of interest, its end points, $T_I = 0$ and $T_I = 25$, define a line segment in the space of the mapping Ω ; this line segment has end points given by the images of $T_I = 0$ and $T_I = 25$, i.e.,

at $T_I = 0$

$$\Omega = \frac{4.71 - j4.6}{29} = 0.162 - j0.159$$

while at $T_I = 25$

$$\Omega = \frac{25(4.6 + j4.71) + (4.71 - j4.6)}{29}$$

$$\Omega = 4.13 + j3.91$$

This line is plotted as $A'B'$ in Fig. 4-17. It is obvious that lines GF, OE and DC map in a similar manner into straight lines in the image (or mapping) space. Also, variations in T_N , for fixed T_L and T_I , map into straight lines; thus, OD, GA, FB and EC can be mapped easily. The results of this mapping yield the end points in the image space for $\omega = 1$ as shown in Table 4-1.

These lines are mapped accordingly, Fig. 4-17; their images are shown by

TABLE 4-1

| Line Mapped | T_L | T_I | T_N | End Points of Image Line |
|-------------|-------|-------|-------|--------------------------|
| AB | 5.3 | 0 | 0.7 | $0.162 - j0.159$ |
| | 5.3 | 25 | 0.7 | $4.13 + j3.91$ |
| GF | 5.3 | 0 | 0 | $0.0345 - j0.183$ |
| | 5.3 | 25 | 0 | $4.61 + j0.679$ |
| OE | 0 | 0 | 0 | 1 |
| | 0 | 25 | 0 | $1 + j25$ |
| DC | 0 | 0 | 0.7 | $1 + j0.7$ |
| | 0 | 25 | 0.7 | $-16.5 + j25.7$ |
| OD | 0 | 0 | 0 | 1 |
| | 0 | 0 | 0.7 | $1 + j0.7$ |
| GA | 5.3 | 0 | 0 | $0.0345 - j0.183$ |
| | 5.3 | 0 | 0.7 | $0.1625 - j0.1586$ |
| FB | 5.3 | 25 | 0 | $4.61 + j0.679$ |
| | 5.3 | 25 | 0.7 | $4.12 + j3.91$ |
| EC | 0 | 25 | 0 | $1 + j25$ |
| | 0 | 25 | 0.7 | $-16.5 + j25.7$ |

the same letter designations but primed. Thus, faces ODCEO and GABFG of Fig. 4-16 map into the triangles and line segment appendages of $O'D'C'E'O'$ and $G'A'B'F'G'$ of Fig. 4-17, respectively.

The mapping of lines AD, GO, FE and BC into Ω space can now be considered. The most general case, i. e., $T_I, T_N \neq 0$, is represented by line BC. Again, for $\omega = 1$,

$$\begin{aligned}\Omega &= \frac{(j25+1)(j0.7+1)}{(T_L j+1)} \\ &= \frac{-16.5+j25.7}{1+jT_L}\end{aligned}\quad (4-13)$$

Equation (4-13) is recognized as the equation of a circle; a proof is given in Appendix A.⁽¹⁾ Since three points determine a circle, the arc corresponding to (4-13) can be found from the two end points on the range of T_L and the property that as $T_L \rightarrow \infty$ the circle passes through the origin, i. e.,

$$\Omega(j\omega) \rightarrow 0 \text{ as } T_L \rightarrow \infty.$$

Thus, on BC the three points are:

$$\begin{array}{ll} T_L \rightarrow \infty & \frac{H_o^n}{H_o}(j1) = 0 \\ T_L = 0 & \frac{H_o^n}{H_o}(j1) = -16.5 + j25.7 \\ T_L = 5.3 & \frac{H_o^n}{H_o}(j1) = 4.13 + j3.9 \end{array}$$

(1) Churchill, R. V., "Complex Variable and Applications", 2nd Ed., McGraw-Hill, New York, 1960.

Alternatively, the equation of the circle is given by (A-8) where in this case, $a = -16.5$, $b = 25.7$ yielding

$$(u + 8.25)^2 + (v - 12.85)^2 = (15.28)^2 \quad (4-14)$$

Of course, only the arc of the circle corresponding to the range $0 \leq T_L \leq 5.3$ is of interest. In a similar manner, the mappings of lines GO, AD and FE are obtained; the information, for $w=1$, is assembled in Table 4-2.

TABLE 4-2

| Line Mapped | T_L | T_I | T_N | End Points of Image Arc | Equation of Corresponding Circle |
|-------------|-------|-------|-------|-------------------------|--------------------------------------|
| BC | 0 | 25 | 0.7 | $-16.5 + j25.7$ | $(u + 8.25)^2 + (v - 12.85)^2$ |
| | 5.3 | 25 | 0.7 | $4.13 + j3.9$ | $= (15.28)^2$ |
| FE | 0 | 25 | 0 | $1 + j25$ | $(u - \frac{1}{2})^2 + (v - 12.5)^2$ |
| | 5.3 | 25 | 0 | $4.61 + j0.679$ | $= (12.53)^2$ |
| GO | 0 | 0 | 0 | 1 | $(u - \frac{1}{2})^2 + v^2$ |
| | 5.3 | 0 | 0 | $0.0345 - j0.183$ | $= (\frac{1}{2})^2$ |
| AD | 0 | 0 | 0.7 | $1 + j0.7$ | $(u - \frac{1}{2})^2 + (v - 0.35)^2$ |
| | 5.3 | 0 | 0.7 | $0.1623 - j0.1586$ | $= (0.6105)^2$ |

It should be observed that, interestingly enough, the equations of the circles which represent variation in T_L are not themselves functions of T_L , although the end points of the arc segments reflecting the range of T_L surely

are. Hence, as long as the ranges on T_I and T_N remain unchanged so do the equations of the circles. This property proves to be extremely useful in this study later on, and will provide an opportunity for obtaining insight and intuitive understanding of the effect of human operator adaptation upon any given system.

The end result of Tables 4-1 and 4-2 is shown plotted in Fig. 4-18 and represents the mapping (3-12) plotted for the single frequency $\omega = 1$. A sufficient number of such plots must be made for selected frequencies to permit determination of whether the locus of (3-11), the difference vector between $-L_n$ and $\frac{H_o}{H_o} = \Omega$, undergoes encirclements of its origin or not. From this, stability information may be inferred. Since the human-operator transfer function is valid only over a limited frequency range, the plots are not infinite in number; additionally, their tendencies are obvious and a few judiciously chosen discrete frequencies are seen to suffice for the purposes of this study. Several points are worth noting at this time:

1. The human operator's pole variations map into straight line segments while his zero variations map into arc segments.
2. The formation of a largest polygon (with some arc boundaries) around the bounds of the region in Fig. 4-18, i. e., formation of the convex hull, defines an area of possible criticality re stability.
3. It must be determined if the end-point mapping results in a complete mapping, i. e., every point of the mapping Ω must be a point of the end-point mapping if the mapping is to be complete.
4. If the hull is not complete, rather than use a convex hull, can the "end point" hull be shown to be almost complete?
5. The situation has changed progressively through the usual Nyquist condition where the critical point is $-1 + j0$, to a one parameter variation where the locus of critical points is a curve, to a final condition where the critical points are defined by an area Ω .

It is shown that Point 4. pertains, but that the hull of the complete mapping can always be found and may be described analytically; however, the end point mapping of Fig. 4-18 differs so little from the complete hull (indeed, very little

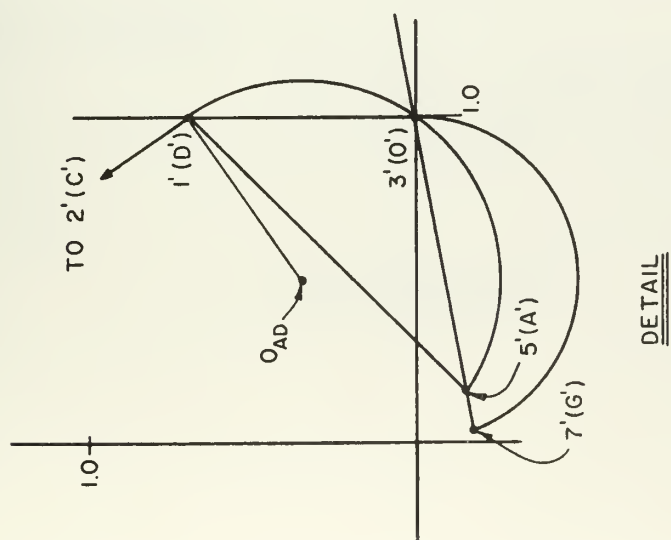
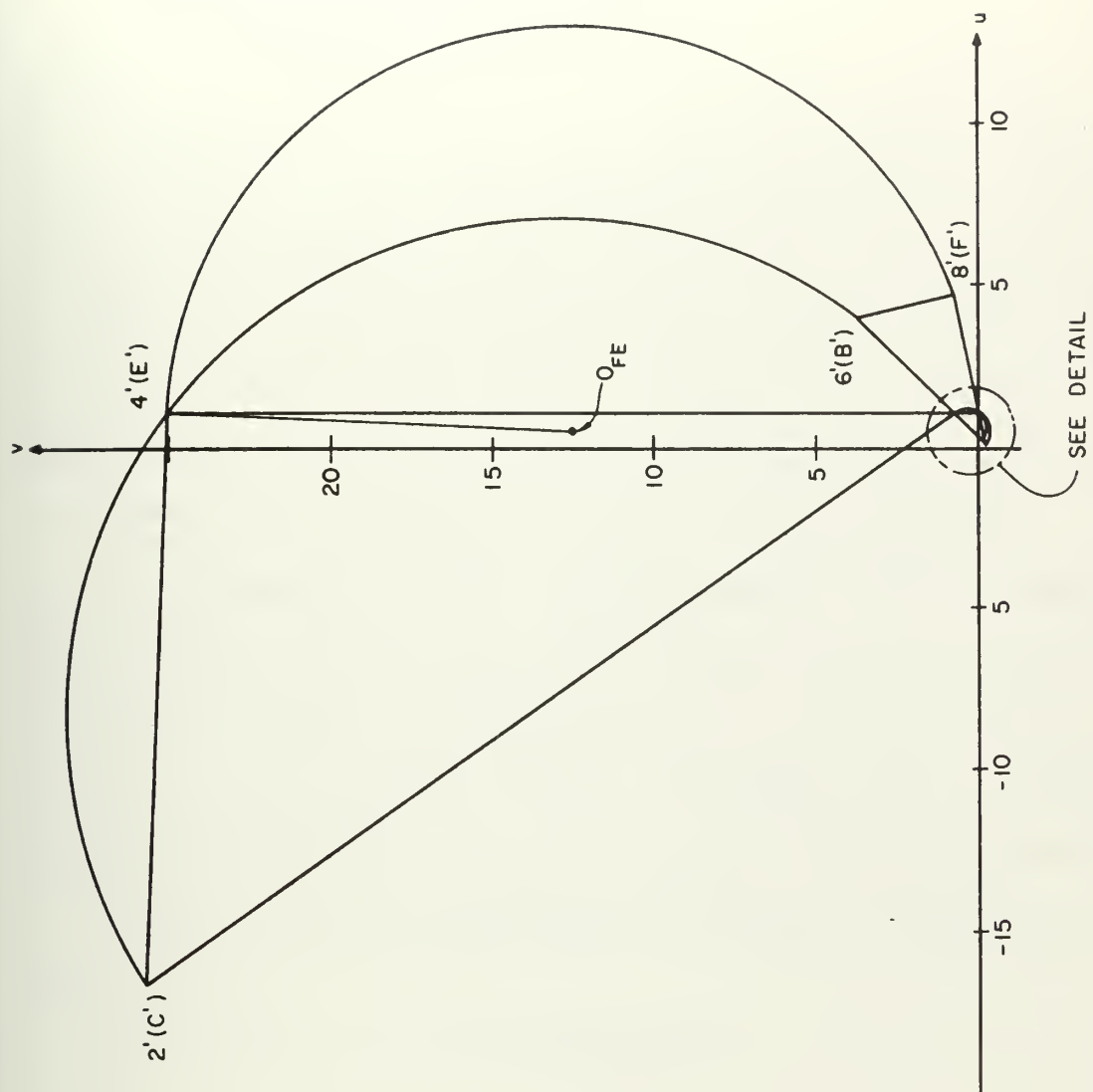


Fig. 4-18 : Typical Human Operator Variable Parameter Mapping ($w=1$)

from the convex hull also) that for all practical purposes it may be used.

The interpretation of Point 5. is: if there is some area of Ω unenclosed by the contour $-L_n$, then presumably the human operator can assume such parameter values and stabilize the system. The parameter values necessary to accomplish this give insight into the amount of effort the human operator must expend to retain control. If the entire Ω region is enclosed, the human operator cannot control under any circumstances (for the capability ranges assumed and plotted). If no portion of the Ω region is enclosed, there is stability no matter what the actions of the human operator (within the sense of the transfer function description). Later, examples and discussion clarify these notions.

For the study to proceed, the regions of varying parameters, Ω , must be plotted for selected frequencies. Generally speaking, the human-operator transfer function has validity only over a range of frequencies, $\omega < 20$. It can be seen that beyond an $\omega = 3$, the plots offer little additional information and their trend is obvious. The region Ω is plotted therefore, for $\omega = 0.5, 1.0, 2.0, 3.0$ and also for an $\omega = 10$; in order to do this it is necessary to derive the mapping relations for the various bounds of Fig. 4-16 as functions of ω . Two such derivations are given below; the first for a human-operator pole and the next for a human-operator zero, the remainder are obtained in a similar way.

The human-operator pole variation, T_I , along AB of Fig. 4-16, is found from

$$\Omega = \frac{(T_I j\omega + 1)(0.7j\omega + 1)}{(5.3j\omega + 1)}$$

After some manipulation,

$$\Omega = \frac{(T_I j\omega + 1)(1 + 3.71\omega^2 - 4.6j\omega)}{1 + 28\omega^2} \quad (4-15)$$

Thus, at $T_I = 0$,

$$\Omega = \frac{1 + 3.71\omega^2 - 4.6j\omega}{1 + 28\omega^2} \quad (4-16)$$

and, at $T_I = 25$,

$$\Omega = \frac{1 + 118.71\omega^2 + j\omega(92.8\omega^2 + 20.4)}{1 + 28\omega^2} \quad (4-17)$$

The mapping of AB is a straight line between the points obtained from (4-16) and (4-17). For variation in the human-operator zero, T_L , along BC in Fig. 4-16,

$$\begin{aligned} \Omega &= \frac{(25j\omega + 1)(0.7j\omega + 1)}{(T_L j\omega + 1)} \\ &= \frac{(-17.5\omega^2 + 25j\omega + 0.7j\omega + 1)}{1 + T_L j\omega} \end{aligned}$$

If the complex number $\Omega(j\omega)$, is represented by $u + jv$, then

$$\begin{aligned} 1 + T_L j\omega &= \frac{1 - 17.5\omega^2 + 25.7j\omega}{u + jv} \cdot \frac{u - jv}{u - jv} \\ &= \frac{u - 17.5\omega^2 u + 25.7j\omega u - jv + 17.5j\omega^2 v + 25.7\omega v}{u^2 + v^2} \end{aligned}$$

The real parts may be set equal to each other (the imaginary parts permit recovery of T_L for given ω , Appendix A):

$$1 = \frac{u - 17.5\omega^2 u + 25.7\omega v}{u^2 + v^2}$$

If the square is completed, one obtains

$$\left[u - \left(\frac{1 - 17.5\omega^2}{2} \right) \right]^2 + \left[v - \frac{25.7\omega}{2} \right]^2 = \left(\frac{1 - 17.5\omega^2}{2} \right)^2 + \left(\frac{25.7\omega}{2} \right)^2 \quad (4-18)$$

Relation (4-18) is the equation of a circle, an arc segment of which, BC maps into, as a function of frequency.

At $T_L = 0$:

$$= 1 - 17.5s^2 + 25.7j \quad (4-19)$$

and at $T_L = 5.3$:

$$= \frac{1 + 118.6s^2 + j(92.8s^2 + 20.4)}{1 + 28s^2} \quad (4-20)$$

In a similar manner, corresponding relations are obtained for the remaining bounds of Fig. 4-16; these are displayed in Table 4-3.

It is now possible to plot the mappings of the parameter variations for particular frequencies; these plots are shown for $\omega = 0.5, 1.0, 2.0, 3.0$ and 10.0 in Figs. 4-19, 4-20, 4-21, 4-22 and 4-23, respectively. The question of the completeness of these end point mappings may now be considered. As is clear from Fig. 4-24 the end point mapping is not significantly different from the convex hull. This is particularly true in the light of system plots of $-L_n$ contours for typical physical plants, as can be seen in later examples. The difference between the end point hull and the true hull (complete mapping), however, is even less than between the end point hull and the convex hull. In the complete mapping, a small parabolic arc extends between segment $O'D'$ and $D'C'$, as well as between $G'A'$ and $A'B'$. For the purposes of this study, this situation may be ignored. Details and a proof are contained in Appendix B.

One further point of some significance relating to the mapping should be noted. As stated above, face $ODCEO$ of Fig. 4-16 maps into the triangle plus line segment $O'D'C'E'O'$, shown in Fig. 4-17; similarly, face $GABFG$ maps into $G'A'B'F'G'$. The former is associated with a $T_L = 0$, while the latter corresponds to a $T_L = 5.3$; this value represents the upper bound on the range for T_L used in this study. It is interesting to observe, and useful later in the physical interpretations, that as T_L varies from 0 to 5.3, face $O'D'C'E'O'$ "rotates" into coincidence with face $G'A'B'F'G'$; the "rotation" occurs by virtue of constrained motion of the significant apex points of the "triangle" along the circular arc segments of Fig. 4-18.

It may be preferable to reflect upon the effects of a two parameter mapping

TABLE 4-3

| Line Mapped | End Point Value of Varying Parameter | General Equations of the curve |
|-------------|---|--|
| AB | $T_I = 0$ | $\frac{1 + 3.71\omega^2 - 4.6j\omega}{1 + 28\omega^2}$ |
| | $T_I = 25$ | $\frac{1 + 118.71\omega^2 + j\omega(92.8\omega^2 + 20.4)}{1 + 28\omega^2}$ |
| GF | $T_I = 0$ | $\frac{1 - 5.3j\omega}{1 + 28\omega^2}$ |
| | $T_I = 25$ | $\frac{1 + 132.5\omega^2 + 19.7j\omega}{1 + 28\omega^2}$ |
| OE | $T_I = 0$ | 1 |
| | $T_I = 25$ | $1 + 25j\omega$ |
| DC | $T_I = 0$ | $1 + .7j\omega$ |
| | $T_I = 25$ | $1 - 17.5\omega^2 + 25.7j\omega$ |
| FB | $T_N = 0$ | $\frac{1 + 132.5\omega^2 + 19.7j\omega}{1 + 28\omega^2}$ |
| | $T_N = 0.7$ | $\frac{1 + 118.71\omega^2 + j\omega(92.8\omega^2 + 20.4)}{1 + 28\omega^2}$ |
| OD | $T_N = 0$ | 1 |
| | $T_N = 0.7$ | $1 + .7j\omega$ |

TABLE 4-3 (cont'd.)

| Line Mapped | End Point Value of Varying Parameter | General Equations of the curve |
|-------------|---|---|
| GA | $T_N = 0$ | $\frac{1 - 5.3j\omega}{1 + 28\omega^2}$ |
| | 0.7 | $\frac{1 + 3.71\omega^2 - 4.6j\omega}{1 + 28\omega^2}$ |
| EC | $T_N = 0$ | $\frac{1 + 25j\omega}{1 + 25j\omega}$ |
| | 0.7 | $\frac{1 - 17.5\omega^2 + 25.7j\omega}{1 + 25j\omega}$ |
| BC | $T_L = 0$ | $\frac{1 - 17.5\omega^2 + 25.7j\omega}{1 + 25j\omega}$ |
| | $T_L =$ | $\left\{ \begin{aligned} &[u - (.5 - 8.75\omega^2)]^2 + [v - 12.85\omega]^2 \\ &= (.5 - 8.75\omega^2)^2 + (12.85\omega)^2 \end{aligned} \right\}$ |
| | 5.3 | $\frac{1 + 118.71\omega^2 + j\omega(92.8\omega^2 + 20.4)}{1 + 28\omega^2}$ |
| | 0 | 1 |
| GO | $T_L =$ | $\left\{ \left[u - \frac{1}{2} \right]^2 + v^2 = \left[\frac{1}{2} \right]^2 \right\}$ |
| | 5.3 | $\frac{1 - 5.3j\omega}{1 + 28\omega^2}$ |
| | 0 | $\frac{1 + .7j\omega}{1 + .7j\omega}$ |
| | $T_L =$ | $\left\{ \left[u - \frac{1}{2} \right]^2 + [v - .35\omega]^2 = \left(\frac{1}{2} \right)^2 + (.35\omega)^2 \right\}$ |
| AD | $T_L =$ | $\frac{1 + 3.71\omega^2 - 4.6j\omega}{1 + 28\omega^2}$ |
| | 0 | $\frac{1 + 25j\omega}{1 + 25j\omega}$ |
| FE | $T_L =$ | $\left\{ \left[u - \frac{1}{2} \right]^2 + [v - 12.5\omega]^2 = \left(\frac{1}{2} \right)^2 + (12.5\omega)^2 \right\}$ |
| | 5.3 | $\frac{1 + 132.5\omega^2 + 19.7j\omega}{1 + 28\omega^2}$ |

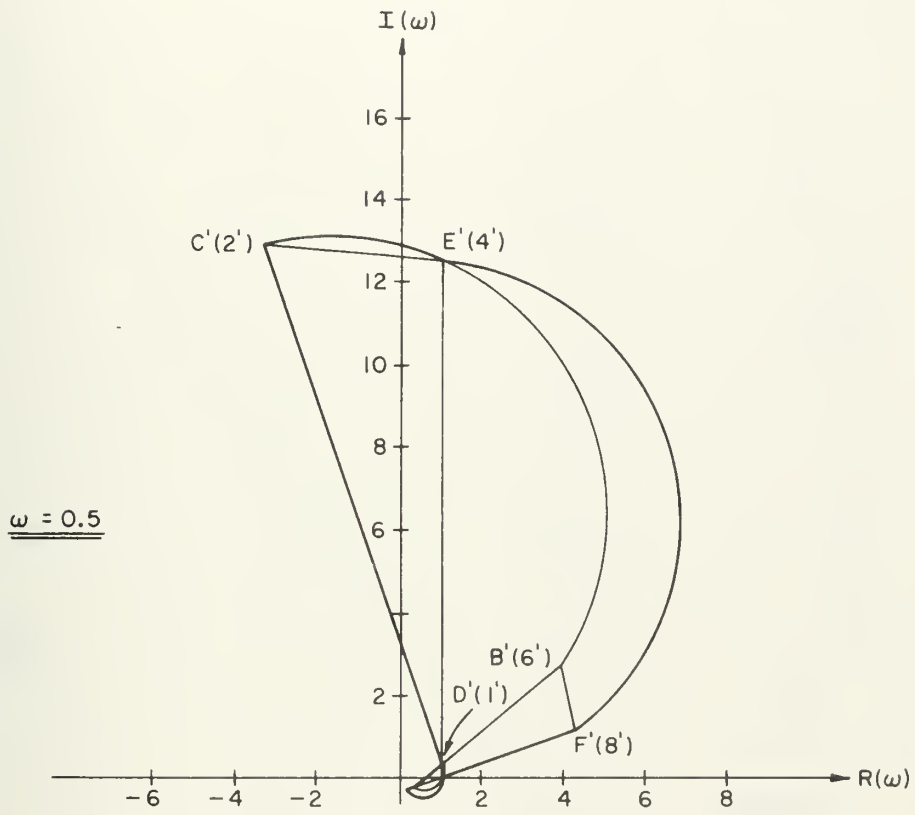


Fig. 4-19 : Variable Parameter Mapping

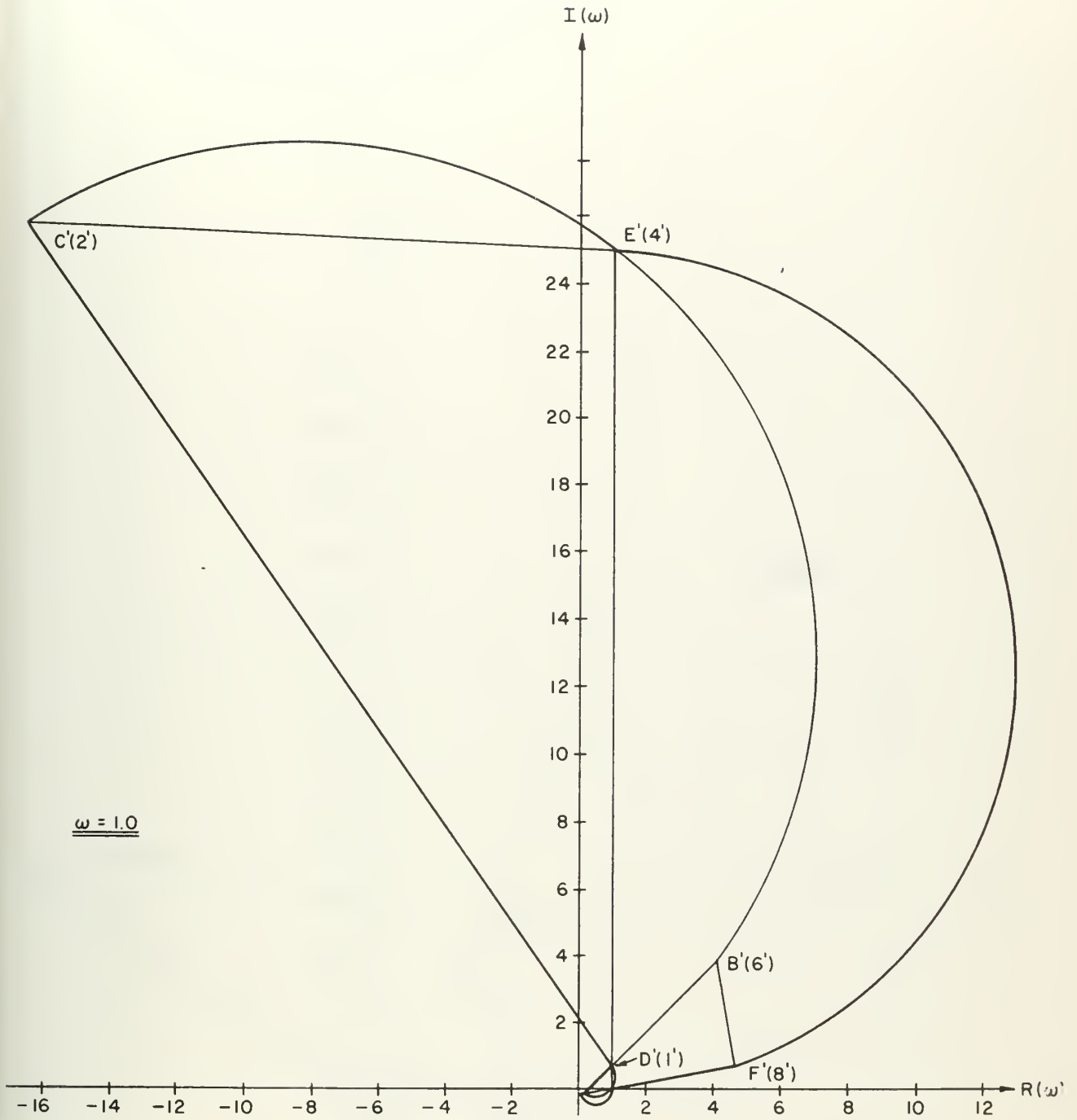


Fig.4-20 : Variable Parameter Mapping

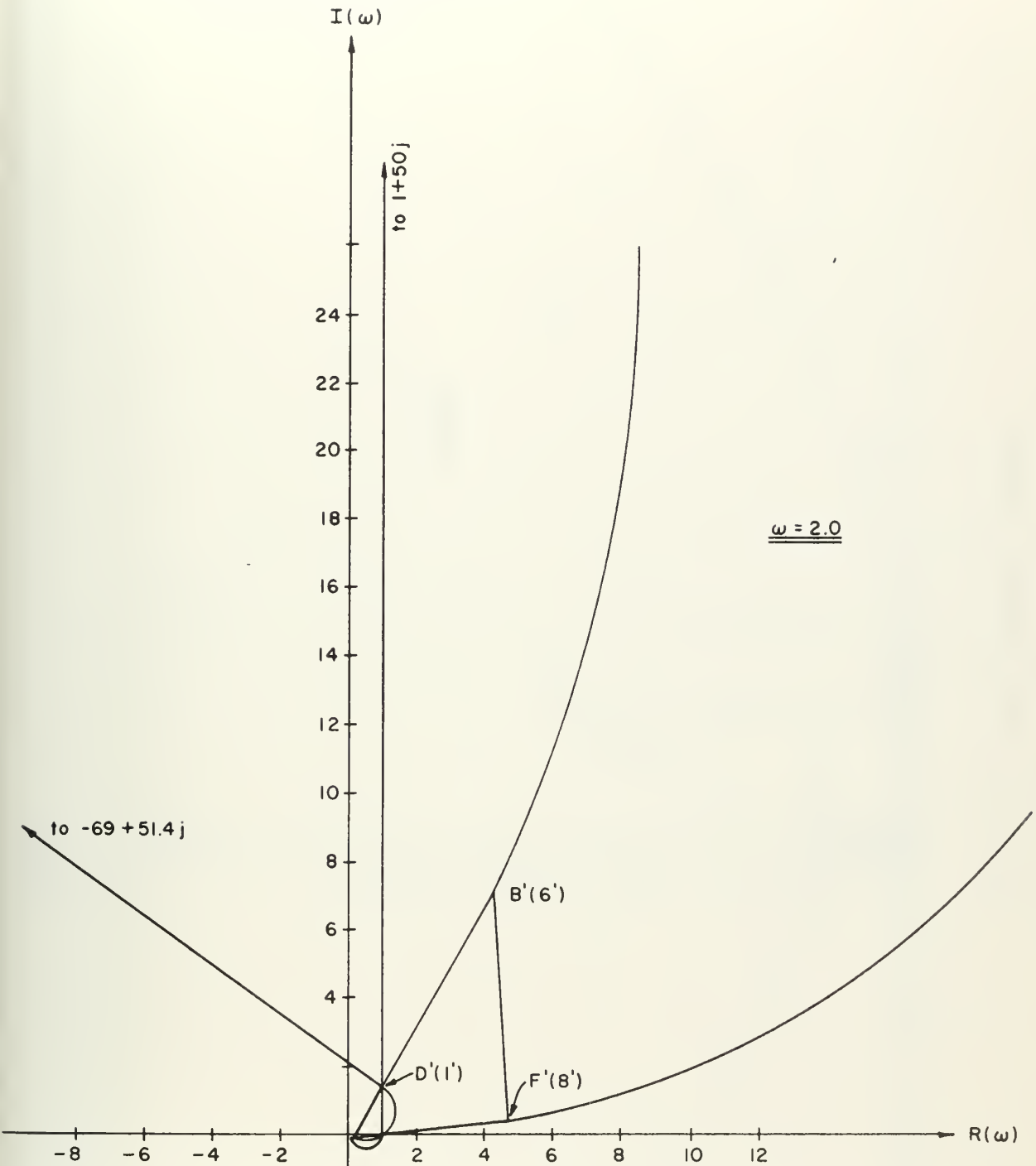


Fig. 4-21 : Variable Parameter Mapping

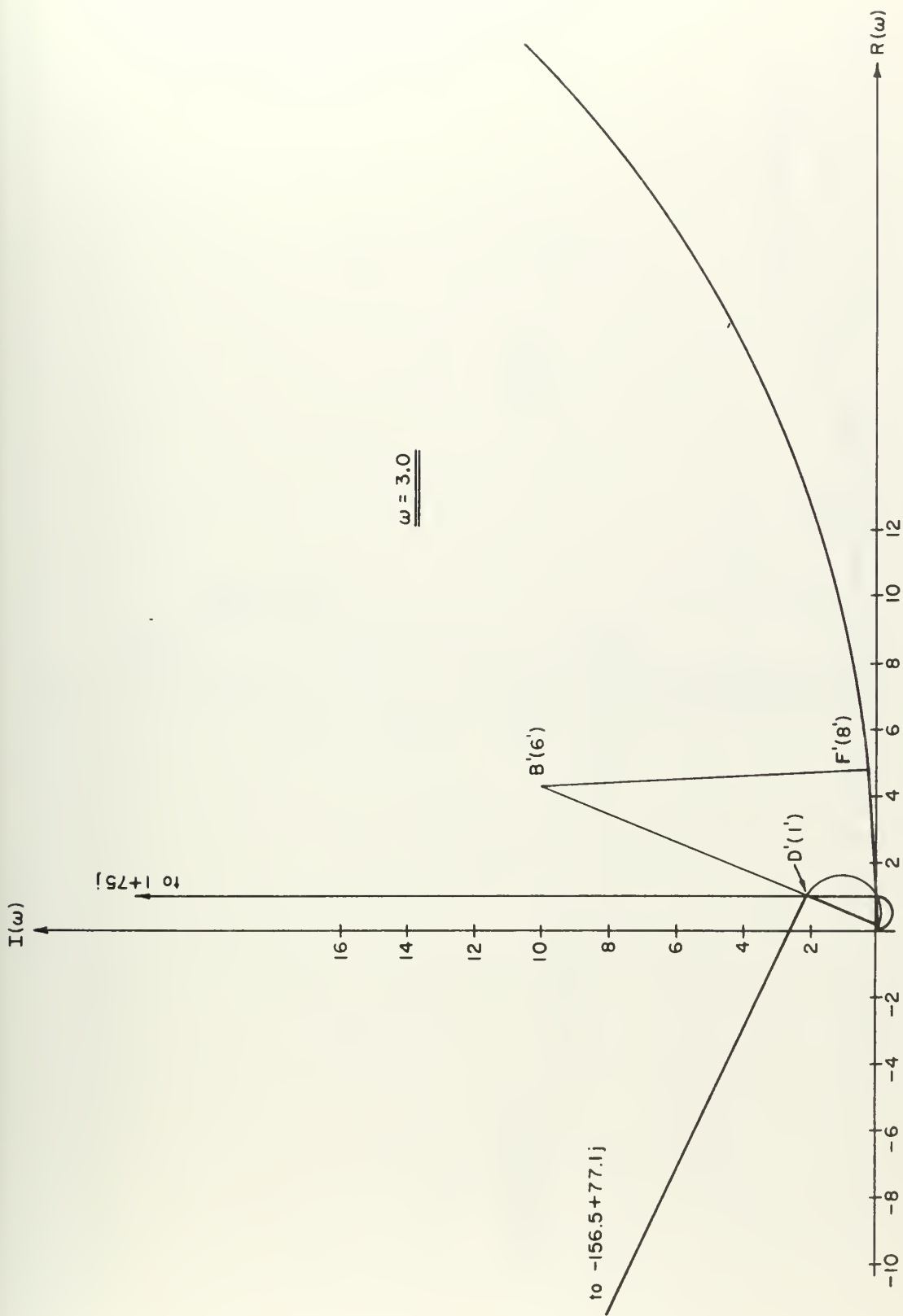


Fig. 4-22: Variable Parameter Mapping

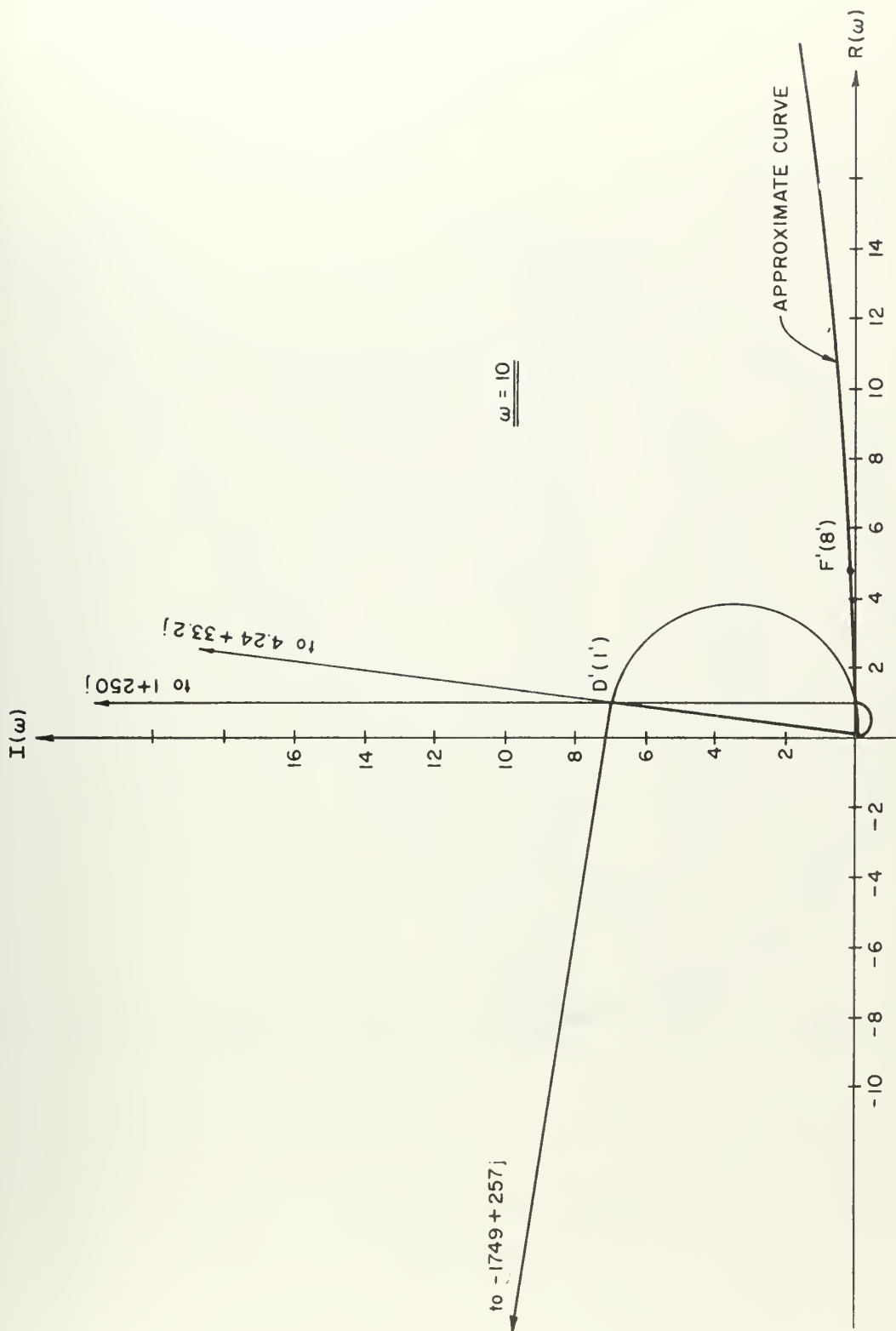


Fig. 4-23 : Variable Parameter Mapping

before continuing on to the practical results, derived from the three parameter mapping, which are described in the next chapter. To this end, and also to offer further stability interpretation practice, consideration of Appendix C may be appropriate at this time.

4.5 Summary

After the analysis of a simple system containing only one variable parameter, the situation of multi-parameter variation is considered. With the familiarity gained from the single parameter example the generalization to a multiple parameter problem follows more easily. In particular, the three-varying parameter problem (one zero and two poles) is considered in detail, since it relates directly to the three time constants in the human-operator transfer function that are within the adaptive capacity of the human controller. It is shown that these parameters, forming a type of variable structure, can be used as a mapping function in the complex plane to display the boundaries of human-adaptive capacity. These boundaries contain a collection of "critical points", akin to the $-1 + j0$ point on the usual Nyquist diagram; when the plant Nyquist contour is plotted in conjunction with these bounded areas (at particular frequencies) interpretation of system stability characteristics is possible (Chapter Five). Several plots of such boundaries are presented for important frequencies in the human range.

CHAPTER FIVE

APPLICATIONS OF THE THEORY5.1 Introduction

The groundwork which is laid in Chapters Two, Three, and Four now makes it possible to consider actual physical systems which are operated by human controllers. If the empirical results can be predicted (or implied) by the approach outlined in the previous chapters, then this technique for viewing the response characteristics of systems with human operators in the loop is justified, and difficulties with human controllability of such systems can be anticipated via the theory.

Of the several experimental studies performed with human controllers, two are particularly well suited to the purposes of this investigation. The Hall study⁽¹⁾ is notable for its comprehensive nature and for the use of a non-trivial controlled element. The predictions obtained from the theory presented in the previous chapters can be checked for their physical validity using the results of the Hall experiment. The Smith data⁽²⁾ is derived from human attempts to control an unstable system; with these empirical results, the effect of open-loop instability upon the theory also may be compared with actual physical accomplishment.

5.2 The Hall and Smith Studies

In 1956 and 1957 a simulator study of the human pilot was conducted by Ian Hall at Princeton University.⁽¹⁾ The intent was to investigate the effects upon the pilot of various controlled-element-longitudinal dynamics. A very excellent analysis by McRuer et al, uses the data of that report as a basis for developing

(1) Hall, I. A. M., Effects of Controlled Element on the Human Pilot, Princeton Univ., WADC Tech. Rept. No. 57-509, Aug., 1958.

(2) Smith, R. H., An Experimental Determination of the Limits of Pilot Controllability for Unstable, Oscillatory Second Order Systems, Thesis for degree of Aeronautical Engineer, Univ. of Cincinnati, 1961 (also appeared as a paper, On the Limits of Manual Control, IEEE Trans. on Hum. Factors in Electronics, Sept., 1963).

a "flying qualities" theory.⁽¹⁾ Briefly, the McRuer approach follows the "adjustment rules" procedure outlined in Section 2.1 of this study. Another experimental study conducted at the University of Cincinnati by Ralph Smith probed for the human limits in controlling certain open-loop-unstable systems.⁽²⁾ The results of both these studies are used to significant advantage here; indeed, the empirical evidence of Smith⁽²⁾ and Hall⁽³⁾ furnished the experimental justification for the results of this study.

5.3 The Hall Study

The experiments were conducted in a mockup cockpit simulating a Navion aircraft. The subjects were two Navy pilots, one with considerable jet time (Pilot I) and the other with a preponderance of propeller aircraft experience (Pilot II)*. The seemingly subjective pilot rating scale of the Hall study used the following definitions (quoted):⁽³⁾

"GOOD (G): aircraft response very suitable for the given tracking task.

ACCEPTABLE (A): possessed of satisfactory flying qualities, but less easy to control than a "good" configuration; aircraft response either faster or slower than that which the pilot visualizes as most desirable.

POOR (P): possessing characteristics which require considerable concentration to accomplish the given tracking task; configurations are either quite touchy, unresponsive, or poorly damped.

UNACCEPTABLE (U): quite unsuitable.

* The author discussed the study with Pilot II in November, 1964.

(1) McRuer, D. T., I. L. Ashkenas, C. L. Guerre, A Systems Analysis View of Longitudinal Flying Qualities, WADD Tech. Rept. 60-43, Wright Air Dvlpmt. Div., Jan., 1960.

(2) Smith, R. H., An Experimental Determination of the Limits of Pilot Controllability for Unstable, Oscillatory Second Order Systems, Thesis for degree of Aeronautical Engineer, Univ. of Cincinnati, 1961 (also appeared as a paper, On the Limits of Manual Control, IEEE Trans. on Hum. Factors in Electronics, Sept., 1963.)

(3) Hall, I. A. M., Effects of Controlled Element on the Human Pilot, Princeton Univ., WADC Tech. Rept. No. 57-509, Aug., 1958.

The subjects differed only between P and U. Pilot II tended to classify aircraft which were unsuitable for the given tracking task, but could conceivably be flown, as P, reserving U for configurations which he considered as unsafe to fly under any circumstances, while Pilot I tended to classify all configurations with which tracking was very difficult as U." Due to the use of plus, "somewhat more than", and minus, "somewhat less than" modifiers on the G, A and P ratings the resultant number of classifications was ten, i.e., $A + \rightarrow 4$, $G \rightarrow 2$, $G + \rightarrow 1$, etc. Those familiar with the military aircraft rating scheme are aware that the standard Cooper scale also has ten gradations; they are not exactly the same.⁽¹⁾ It may appear that the entire rating system is seriously suspect by virtue of its apparent subjectivism. The consistency of pilot opinion, however, is remarkably good and the variation between pilots astonishingly small.⁽²⁾ The inference, of course, is that objective factors are coming to play and it is there that the prospect of a possible quantitative theory for handling qualities arises.

The experiment was conducted in the compensatory arrangement as shown in Fig. 2-1 with the plant (aircraft) dynamics having the generic form (short-period approximation):

$$G_P = \frac{5(1 + 0.6s)\omega_n^2}{s(s^2 + 2\zeta\omega_n s + \omega_n^2)} \quad (5-1)$$

Hall used twenty combinations of ζ and ω_n to test pilot response. Of interest here is a representative sample covering the spectrum of pilot opinion. The configurations designated by Hall as C30, C25, C32, C31, and C21 are to be studied, Table 5-1.

The following pertinent performance comment is a quote from the Hall report.⁽³⁾ "As a measure of tracking efficiency... an average absolute value

(1) Cooper, G.E., Understanding and Interpreting Pilot Opinion, Aero. Eng. Rev. Vol. 16, No. 3., Mar., 1957.

(2) Ashkenas, I. L. and D. T. McRuer, A Theory of Handling Qualities Derived from Pilot-Vehicle System Considerations, Aerospace Engineering, Feb., 1962.

(3) Hall, I. A. M., Effects of Controlled Element on the Human Pilot, Princeton Univ., WADC Tech. Rept. No. 57-509, Aug., 1958.

TABLE 5-1

| Configuration | | | Opinion Rating | | | Mean Absolute Tracking Error $ \bar{\epsilon} $ (degrees) | | |
|---------------|------------|---------|----------------|----------|---------|---|----------|---------|
| C number | ω_n | ζ | Pilot I | Pilot II | Average | Pilot I | Pilot II | Average |
| C30 | 2.51 | 1.0 | 3 | 1 | 2 | .85 | .64 | .75 |
| C25 | 1.57 | 1.0 | 6 | 3 | 4.5 | .90 | .77 | .84 |
| C32 | 3.77 | 0.35 | 4 | 6 | 5.0 | .83 | .77 | .80 |
| C31 | 3.77 | 0.2 | 5.5 | 8 | 6.75 | .84 | .82 | .83 |
| C21 | 1.57 | 0.2 | 10 | 10 | 10 | 1.30 | 1.04 | 1.17 |

criterion was preferred to a root mean square error, since it was believed that the occasional large errors, which essentially decide the value of RMS error, more often were due to lapses in concentration rather than conditions associated with longitudinal dynamic characteristics... In discussing average error it is interesting to keep in mind that without any pilot output this average would be 1.36 degrees... due to longitudinal... excitation signals... Average longitudinal errors of less than 1.36 degrees indicate that the pilot is doing some good—something which must not be taken for granted."

Additionally, the system excitation was random appearing with Gaussian amplitude probability distribution which resulted in a realistic target course; specifically, the input was white noise cut off at 18db/oct after 1 rad/sec. It is noteworthy that "attempts to increase the cutoff frequency to 2 rad/sec. and 4 rad/sec. had resulted in loss of the pilot's faith in his ability to ever catch the target."⁽¹⁾

In the case of C21, a Routh test shows this configuration to be closed-loop unstable (without a human controller in the loop).

⁽¹⁾ Hall, I.A. M., Effects of Controlled Element on the Human Pilot, Princeton Univ., WADC Tech. Rept. No. 57-509, Aug., 1958.

$$1 + F(s) = 1 + \frac{5\omega_n^2(1 + 0.6s)}{s(s^2 + 2\zeta\omega_n s + \omega_n^2)}; (\zeta = 0.2, \omega_n = 1.57)$$

$$\therefore D(s) = s^3 + 0.628s^2 + 9.84s + 12.30 = 0$$

and

| | | |
|-------|---|-------|
| s^3 | 1 | 9.84 |
| s^2 | .628 | 12.30 |
| s^1 | $\frac{-(12.30 - 6.18)}{.628} \rightarrow 6.12$ | |
| s^0 | $\frac{-(6.12)(12.3)}{(-)} \rightarrow +72$ | |

the two changes of sign in the first column signify two positive roots and therefore closed-loop-system instability. The human operator's task is clear if any attempt is made to control this system.

Next, the closed-loop-stability characteristics of C32 in the Hall report are determined.

$$1 + F(s) = 1 + \frac{5\omega_n^2(1 + 0.6s)}{s(s^2 + 2\zeta\omega_n s + \omega_n^2)}; (\zeta = 0.35, \omega_n = 3.77)$$

$$\therefore D(s) = s^3 + 2.64s^2 + 56.8s + 71 = 0$$

The Routh array is given by

| | | |
|-------|--|------|
| s^3 | 1 | 56.8 |
| s^2 | 2.64 | 71 |
| s^1 | $\frac{-(71 - 150)}{2.64} \rightarrow +89$ | |
| s^0 | $-(-89)(71) \rightarrow +$ | |

since there are no sign changes, there are no positive roots, and the system is

closed-loop stable (without the human operator). It should be expected that less human effort is needed to control this system.

In plotting these systems in conjunction with the variable parameter mapping representing the human adaptive range, one hopes to see the $-L_n(j\omega)$ contour for C21 cross the Ω region in a less acceptable manner than the plot of $-L_n(j\omega)$ for C32 crosses it. The term "a less acceptable manner" may not be adequately defined at this point, but the two extremes of stability are clear, i. e., the case of stability regardless of the pilot's presence, or the other extreme of instability irrespective of the pilots attempts.

The configuration C21 and C32 are plotted in Figs. 5-1 and 5-2, respectively. The solid line through the specified frequency points is merely a suggested locus and is offered mostly for visual ease. As indicated in Chapter Four, the important properties of each system are interpreted from the characteristics of the difference vector drawn between all frequencies of the system locus, $-L_n$, and corresponding frequencies in the variable parameter mapping, Ω . The difficulty, of course, is just what sub-region, set of points, or point is used in the large expanse of Ω ? It is recognized that if a point in the space of the Ω mapping is selected this is equivalent to selecting a set of values for the triple (T_L, T_I, T_N) . It might seem that this amounts to using a specific set of parameter values and working with a nominal transfer function, so the human adaptive range is not really used in the study. This is rather a pedantic point of view. It is seen that all the possible capabilities available to the human (by way of parameter variation) are clearly displayed in the plots and the net effect of choosing any set of them (a triple) is equally visible. The usefulness of such information is undeniable, particularly when experience and experiment allow correlation of certain of the sub-regions of the Ω space with human effort. Such implication is suggested even by the few examples viewed here. In any case, selecting a set of parameter values corresponding to a point in Ω space, or even a cluster of parameter values (T_L, T_I, T_N) , each over a tiny range whose image yields a sub-region in Ω space, is really no more than the human operator is forced to do. The human inability, or reluctance, to change parameter character (break frequencies) discontinuously, corresponds to a constraint upon jumping about in Ω

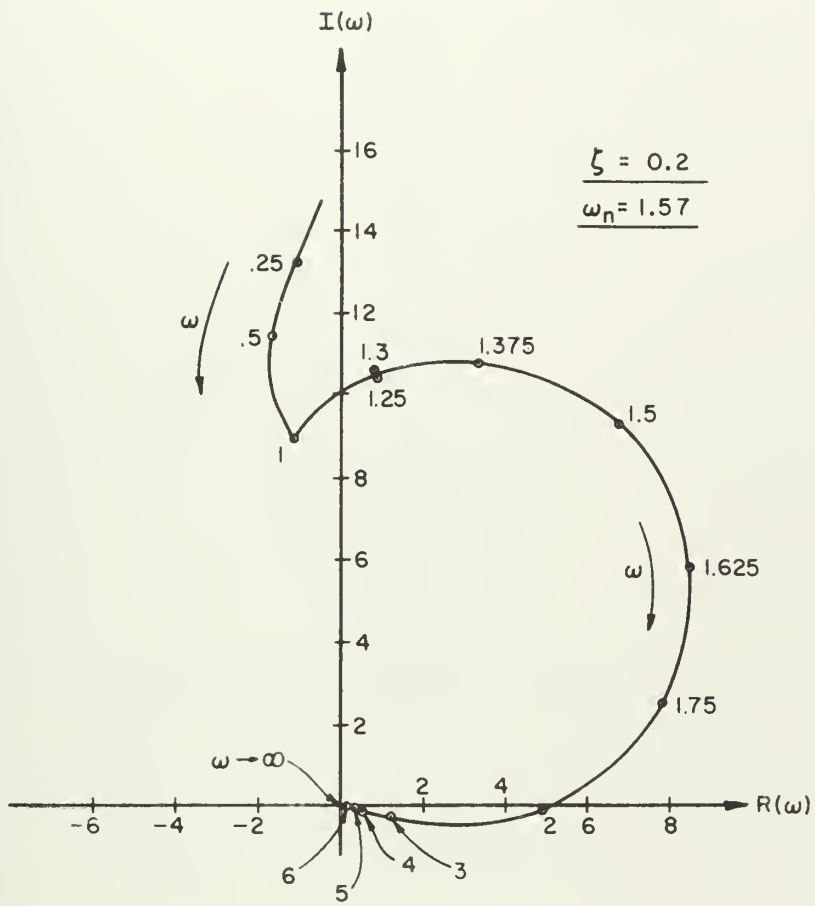


Fig. 5-1: Plot of Configuration C21

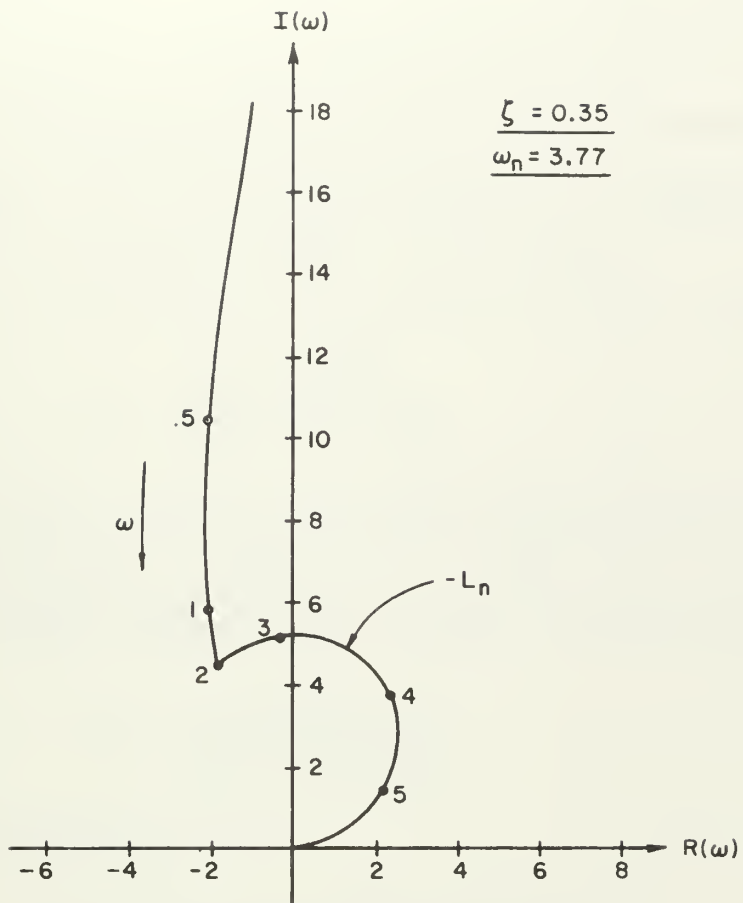


Fig. 5-2 : Plot of Configuration C32

space once some "neighborhood" or sub-region of that space has been chosen. As to how such a neighborhood or sub-region is chosen, some insight regarding this devolves from the examples along with the inherent assumption that the human operator will always try to do his "best". It is not unreasonable to associate "doing one's best" with the maintenance of stability, which is the prime requirement of any system. Once in this sub-region it seems entirely proper that significant departures from it do not occur and interpretations as to whether any portions of this sub-region are "enclosed" or not by $-L_n$ are possible.

From (5-1) is obtained the FIXED portion of the system, the negative of

$$L_n = \frac{5(1 + .6s)\omega_n^2}{s(s^2 + 2\zeta\omega_n s + \omega_n^2)} \quad (5-2)$$

(5-2) is plotted to obtain $-L_n$ (180° phase angle is added). Since the delay time, $e^{-\tau s}$, and human gain, K_h , are lumped with (5-2) it should be borne in mind that Fig. 5-1 actually represents a situation where $K_h e^{-\tau s} = 1$, i. e., $\tau = 0$, $K_h = 1$.

The significance of these conditions comes to focus shortly. Equation (5-2) may be written

$$-L_n = \frac{-5(1 + 0.6j\omega)\omega_n^2}{j\omega(\omega_n^2 - \omega^2 + 2\zeta\omega_n j\omega)} \quad (5-3)$$

Equation (5-3) is rationalized to give

$$-L_n = \frac{10\zeta\omega_n^3\omega^2(1 + 0.6j\omega) + j5\omega_n^2\omega(\omega_n^2 - \omega^2)(1 + 0.6j\omega)}{4\zeta^2\omega_n^2\omega^4 + \omega_n^2(\omega_n^2 - \omega^2)^2} \quad (5-4)$$

The imaginary part of (5-4) is set equal to zero

$$\omega^2(6\zeta\omega_n - 5) + 5\omega_n^2 = 0$$

which for C21 ($\zeta = 0.2$, $\omega_n = 1.57$) gives a real axis crossing at $\omega = \pm 1.99$. For C32, the solutions are imaginary, reflecting no real-axis intercepts. The conditions are portrayed in Fig. 5-1 and 5-2, respectively.

Transparent overlays of Figs. 4-19 through 4-23 are enclosed in an enve-

lopes fastened to the inside back cover of this study. These overlays, which are universally applicable, for the given scale, facilitate considerably the determination of system stability characteristics. Detailed consideration of configuration C21 follows.

A glance at Fig. 5-1 and the general trend of the Ω mappings for the frequencies given ($\omega = 0.5, 1.0, 2.0, 3.0, 10.0$), indicates that system stability is determined by the relation between the $-L_n$ plot and the overlays for $\omega = 2.0$ and 3.0 . This is obvious, since the $-L_n$ plot leaves the Ω region of criticality for $\omega \approx 2$, thus implying that a condition of instability may pertain. It is not yet clear how to take the difference vector between $-L_n(j\omega)$ and $\Omega(j\omega)$, i. e., it could conceivably be connected from $-L_n(j\omega)$ to any point within or on $\Omega(j\omega)$, and for instability it must be demonstrated that a clockwise encirclement must occur as $\omega \rightarrow \infty$, irrespective of this vector's selected orientation (all the Hall configurations were open-loop stable). Due to the time invariance assumptions, however, once a settled condition of operation has been adopted by the human controller, any "selected orientation" of the difference vector defines the sub-region in space to be used for the remainder of the stability analysis. One notes, for example, that if the human operator assumes a set of parameters near point B' of Fig. 4-21 (corresponding to a neighborhood of $\max T_L$, $\max T_I$, $\max T_N$), there results, from Fig. 5-1, an encirclement by the difference vector and instability. If the point F' is assumed, the human can "hold out" longer before enclosure, but eventually an encirclement again occurs. Unfortunately, the prospects of being at F' are not encouraging since, although the point corresponds to $\max T_L$ (lead) which is obviously needed, it also implies $T_N = 0$ which, giving some credence to the physiological correlations, corresponds to no neuromuscular lag. It is soon discovered that if there is any hope for system stabilization or controllability, the human operator must assume parameter values somewhere between arc $G'O'$ and arc $A'D'$ and in the vicinity of the origin of Fig. 4-21, also Fig. 4-18 (detail). Particularly is this the case when the human time-delay is imposed upon the system. Thus, the characteristic of C21 projecting beyond the Ω region is further aggravated by the introduction of the time lag, i. e., the point $-L_n(j\omega)$, at approximately 5.0 on the real axis, rotates with that distance as a radius arm, a number

of degrees proportional to the time delay. For example, if $\tau = .15$ is assumed, the angle of $e^{-j\omega\tau}$ (its magnitude, of course, is unity) for $\omega=2$ is $(2)(.15) = .3$ radian $= 17.2^\circ$, taken in a clockwise direction (lag), and similarly for other pertinent frequencies; clearly, conditions are not made better. It might be added at this point that the effect of the $e^{-j\omega\tau}$ term should not be considered as adding ever increasing phase, thus further complicating the plot; since the human-operator transfer function is not valid beyond a maximum frequency of $\omega \approx 20$, the corresponding phase angle contributed by this frequency can be regarded as an upper bound.

Since C21 is more than a hypothetical situation, it is possible to check the implications of the above analysis. A "best fit" human-operator description was calculated by Hall from the empirical data and a plot of this function can be made. The Hall fitted transfer function was (Pilot II):

$$H_o = \frac{0.1(1 + \frac{s}{2})^2}{(1 + \frac{s}{6.3})^2} e^{-.15s} \quad (5-5)$$

Several properties of (5-5) should be evident. First, this is one of the cases where Hall required a double lead; this characteristic is not accommodated by the generalized human-operator transfer function assumed for this investigation, Equation (2-1). Nevertheless, the description is analyzed in order that the curve representing the Hall fitted transfer function can be observed relative to the regions of criticality which have been used in this study. To be noted are the value of human gain assumed, i. e., $K_h = 0.1$, and time delay tolerated, $\tau = .15$; the system dynamics, $-L_n$, are plotted for $K_h = 1$ although here it is KNOWN that K_h actually was 0.1. The effect of this gain reduction must be imposed upon the system dynamics as indicated previously; in an actual study, where the approach is from the direction of synthesis, the plots imply that reduced human gain is a necessity for this man-machine system if there is to be any chance of controllability.

Equation (5-5) represents those characteristics assumed by a real life human operator and when mapped properly, indicates those points in Ω space

actually assumed under the task conditions imposed by C21. Table 5-2 displays the pertinent information for

$$\Omega = \frac{(1 + 0.159j\omega)^2}{(1 + 0.5j\omega)^2} \quad (5-6)$$

TABLE 5-2 (C21)

| ω | 0 | .5 | 1 | 2 | 3 | 7 | 10 |
|-------------------|-----|-----------|-----------|-----------|-----------|-----------|-----------|
| $\Omega(j\omega)$ | 1/0 | .94/-18.8 | .82/-34.8 | .55/-54.6 | .39/-61.6 | .17/-52.6 | .01/-41.8 |

The last row in Table 5-2 displays the values of Hall's "best fit" transfer function for several pertinent frequencies. A plot is not required to see that the human controller did, indeed, assume those characteristics previously implied by the analysis; furthermore, it is obvious that his capacity is being strained severely just to maintain control of the C21 system. That is to say, precious few points in the variable mapping space Ω are not enclosed by the plot of Fig. 5-1, particularly after introduction of the time-delay effects. It is also interesting to observe that the double lead does not impose any gross effect upon the experimentally determined Ω that is not already incorporated in the generalized mapping (3-12). Additionally, it is noted that reduced gain, Fig. 5-1, would ease the instability tendencies considerably.

Since the dynamics of C21 were rated "unsatisfactory" by the pilots in the Hall study, it is interesting to look in detail at an analysis similar to that above for C32, to which a rating of "acceptable" was assigned. In Fig. 5-2, the plot for C32 of the Hall study, it is clear that controllability should not be as difficult a task as presented previously by C21. It should be obvious what the trend is for the plots of Figs. 4-19 through 4-22 (the transparent overlays) as the frequency continues to be increased beyond $\omega = 3$, Fig. 4-23. The plot for $\omega = 3$, Fig. 4-22, is already so large as to preclude displaying the entire hull upon a page. For this reason, only the region of greatest significance is displayed in the plot. Thus, observation of the characteristics of C32, Fig. 5-2, should indicate that the human

controller need not reduce his T_I to almost zero while maximizing his lead, thus operating in the neighborhood of the origin of Fig. 4-22, as was required in C21; but, that there is room for assuming parameter values in Ω that yield a critical sub-region in the neighborhood of $3.5 + j1$, Figs. 5-2 and 4-22. If the mathematical modeling here represents in some small way the actual physical situation that obtains, it is to be expected that the human controller would feel less vitriolic about this system.

For the system dynamics of the Hall report, C32, the "best-fitted" transfer function for the pilot was (Pilot II):

$$H_o = 0.5 \frac{(1 + \frac{s}{1.3})}{(1 + \frac{s}{0.4})} e^{-.15s} \quad (5-7)$$

Again the value of gain actually assumed by the pilot was less than that used in the plot of Fig. 5-2; the influence of delay time $\tau = 0.15$ seconds also is not included; both facts must be incorporated in the final interpretations. The associated mapping function for Ω is

$$\Omega = \frac{(1 + 2.5j\omega)}{(1 + 0.77j\omega)} \quad (5-8)$$

The form of equation (5-7) may seem to imply a $T_N = 0$ which would be physiological nonsense. Frequently, experiments connected with tracking studies and the servomechanism approach have incorporated some of the minor lags into the delay time wherever such a technique did not prejudice the data.⁽¹⁾ The pertinent data for (5-8) is given in Table 5-3.

TABLE 5-3 (C32)

| ω | 0 | .5 | 1 | 2 | 3 | 7 | 10 |
|-------------------|---|---------------|----------------|-------------|--------------|---------------|---------------|
| $\Omega(j\omega)$ | 1 | $1.29 + j.75$ | $1.84 + j1.09$ | $2.59 + j1$ | $2.9 + j.82$ | $3.18 + j.40$ | $3.22 + j.29$ |

(1) McRuer, D. T. and E. S. Krendel, The Human Operator as a Servo System Element, Journ. of Franklin Instit., Vol. 267, Nos. 5 and 6, May and June, 1959.

The results of actual human control shown in Table 5-3, viewed in conjunction with Figs. 5-2 and 4-22, again bear out the previous implications of the analysis. It is seen that the human controller did indeed assume those values for his parameter characteristics which the two plots of Fig. 5-2 and 4-22 appeared to indicate. Since not as large a value of T_L was required to stabilize this system, the human operator did not assume any larger value than needed when the time lag is considered, and since experiment has shown that generating lead (anticipation or derivative error control) is regarded by the operator as a strain, the system C32 was regarded as superior to C21 and rated "acceptable". Again, it is obvious from Fig. 5-2, that reduced gain will tend to stabilize the system, this was recognized empirically by Pilot II in the experiment as demonstrated by the gain actually assumed, $K_h = 0.5$.

The plots for system configurations C30, C25, and C31 of the Hall study are shown in Figs. 5-3, 5-4, and 5-5 respectively. It is noted that the shape of the curve for the particular system being considered, as well as the relative location of its frequency markings, lead to estimates of pilot-derived handling evaluations with orderings equivalent to those appearing in Table 5-1 under the column Opinion Rating (average). A quick check of Figs. 5-3 through 5-5 and their relation to the mappings on the transparent overlays corroborates this point. Thus, C30 is regarded as a quite good "aircraft", with C25 and C31 rated lower, in that order. The significance of these facts and their ramifications as discussed in detail for C21 and C32 follow directly.

Any person conducting research in the field of control systems and the human operator soon encounters a preponderance of studies, published and unpublished, using a nominal value for T_L of approximately 2.0; indeed, many experimenters feel that human-generated lead cannot exceed this figure. It is informative if the effects upon the stability hull of assigning such a bound to the range of T_L are considered in some detail. As in Chapter Four changing the upper bound on the range of T_L has no effect upon the associated circles themselves (in the complex plane), but only upon the extent of the arc segments. Thus, for a $T_L = 2$, one expects to see smaller arc lengths than for a $T_L = 5.3$, which has been used up to now. It is possible to infer already the effect upon the sta-

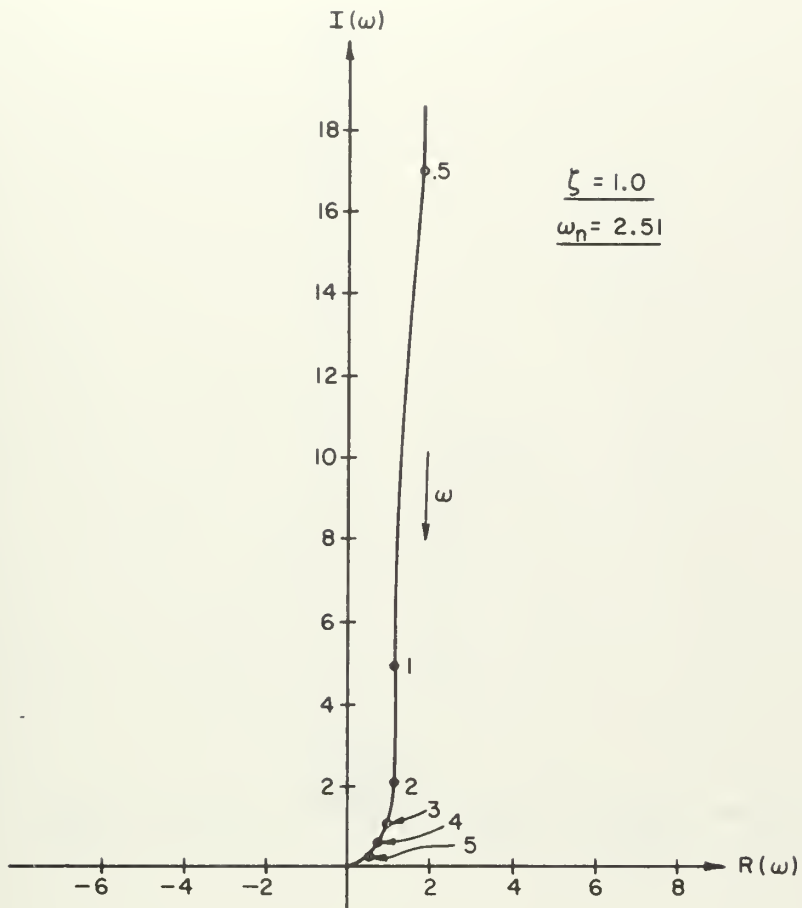


Fig. 5-3 : Plot of Configuration C30

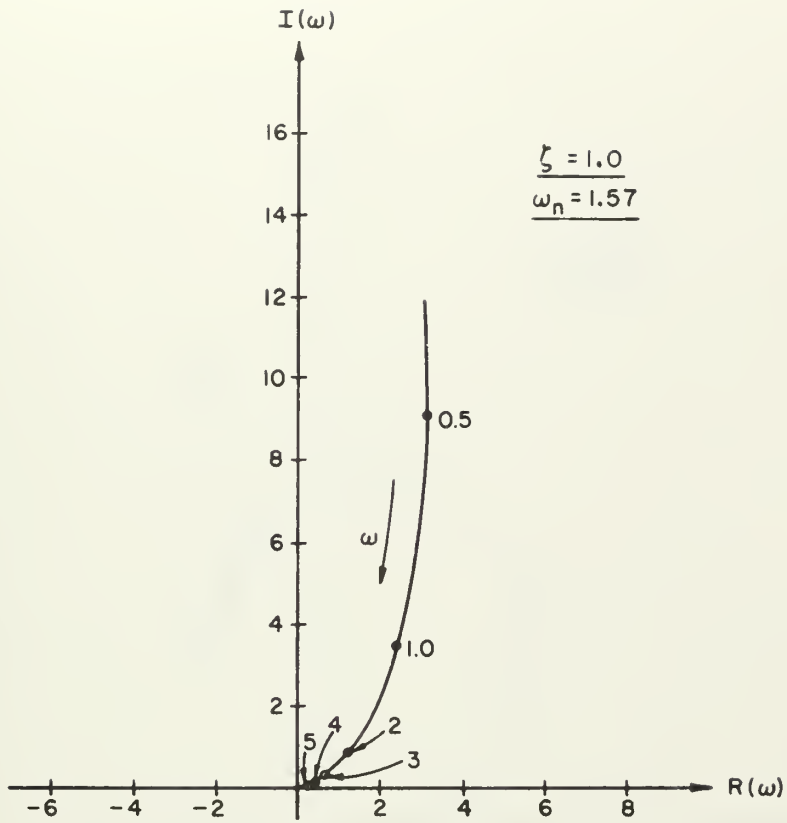


Fig. 5-4 : Plot of Configuration C25

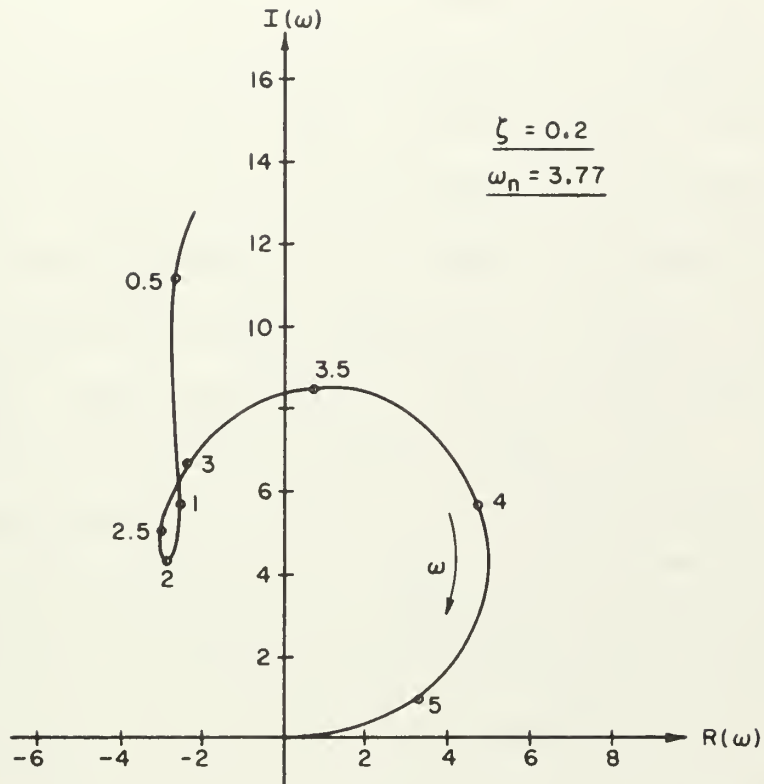


Fig. 5-5 : Plot of Configuration C31

bilizing ability of the human operator without actually plotting the new specifications. One notes that points $5'6'8'7'$ corresponding to face 5687 in three space, Figs. 4-18 and 4-16 respectively, do not move along their circular loci to the extent previously, when $T_L = 5.3$. Thus, line $7'-8'$ makes a steeper angle with the real axis and there are many more systems with amplitude-frequency characteristics which leave the hull and tend toward instability. It is interesting to note that line $7'-8'$ relates to a $T_N = 0$ which, as mentioned above, is usually not possible (unless the effects of T_N can be incorporated into an expanded τ); hence, for $T_N = 0.7$, the stability criterion becomes associated with line $5'-6'$ and further difficulties with stabilization are implied. It is in this way that the designer is given considerable insight into the effects of human-operator parameter variations upon overall system stability and even, perhaps, handling qualities. For example, with maximum lead in, attention may be confined to face 5687 corresponding to the augmented triangular mapping $5'6'8'7'$. Clearly, the outer bound of this "sub-hull" most conducive to stabilizing a system is line $7'-8'$, i. e., relating to max T_L and min T_N . It is plain to see now why all previous examples drawn from empirical results had stability determination dependent upon points clustered in a very reduced area of the large region Ω , i. e., this reduced area corresponded to those extreme values of possible parameter variation most conducive to stability. The consequences of various modifications in human-operator capability are all quickly discerned with this technique, i. e., the region of criticality for stability, if $T_I = 0$, is the mapping of face 1375 which is quickly found to have a hull $1'3'7'5'$ (with two arc segments), and the ramifications of this state of affairs are rapidly seen.* Of course, if the ranges of T_I and T_N must be altered, the hull in the complex plane is changed, but the new points may be found expeditiously from relations such as (4-12) in suitably modified form. The effects of varying T_I and T_N , Appendix B.2, and the characteristic that the T_L circles always pass through the origin, can be

* Appendix C contains an example of the use of the degenerate hull in a two parameter mapping problem as well as an instructive demonstration in stability interpretation.

used to good advantage in plotting any modified range on the parameters T_L , T_I , T_N .

5.4 The Smith Study

In September, 1963 Ralph Smith reported results from an experimental investigation attempting to define the limits of manual control for unstable-second order controlled elements.⁽¹⁾ There are but two results given:

$$[1] \quad \zeta = -.70 \quad \omega_n = 2.5 \quad K_p = 300 \quad (5-9)$$

$$[2] \quad \zeta = -.35 \quad \omega_n = 7.2 \quad K_p = 1200 \quad (5-10)$$

Both of which are for the controlled element having the form

$$G_p = \frac{K_p \omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2} \quad (5-11)$$

To plot Set (5-9), from (5-11) one obtains

$$G_p = \frac{300(6.25)}{s^2 + (-1.4)(2.5)s + 6.25} = \frac{1875}{s^2 - 3.5s + 6.25}$$

For $s = j\omega$, and after some manipulation, it is found that

$$-L_n = \frac{-11718.75 + 1875\omega^2 - j3.5\omega}{(6.25 - \omega^2)^2 + 12.25\omega^2} \quad (5-12)$$

and, in a similar manner for (5-10), but in a non-rationalized form

$$-L_n = \frac{-62.4 \times 10^3}{(51.9 - \omega^2) - j5.04\omega} \quad (5-13)$$

Qualitative plots of (5-12) and (5-13) appear in Figs. 5-6 and 5-7, respectively. The actual coordinate information for (5-12) and (5-13) appears,

⁽¹⁾Smith, R. H., An Experimental Determination of the Limits of Pilot Controllability for Unstable, Oscillatory Second Order Systems, Thesis for degree of Aeronautical Engineer, Univ. of Cincinnati, 1961 (also appeared as a paper, On the Limits of Manual Control, IEEE Trans. on Hum. Factors in Electronics, Sept. 1963.)

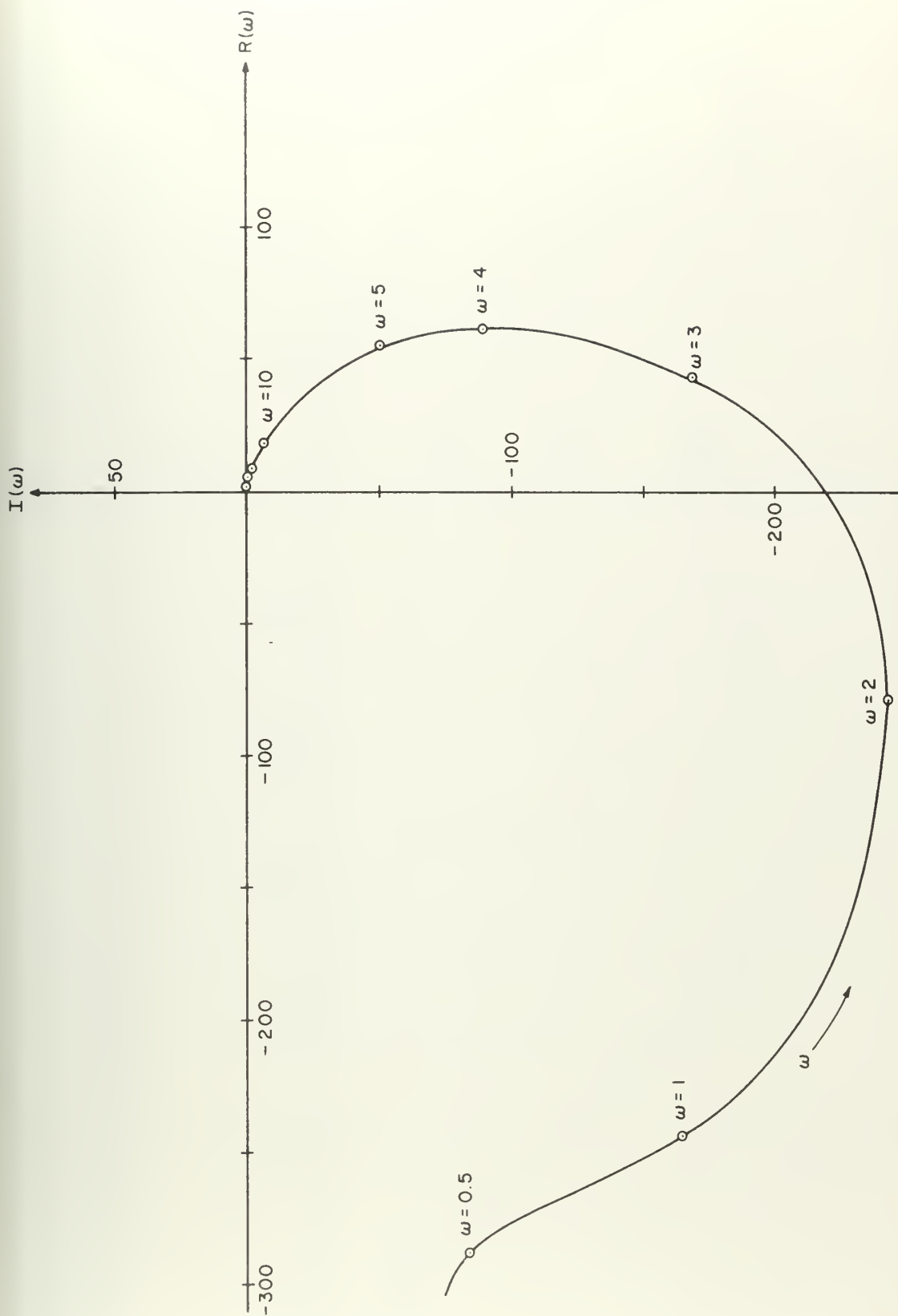


Fig. 5-6 : Open Loop Unstable Plant ($\zeta = -0.7$, $\omega_n = 2.5$)

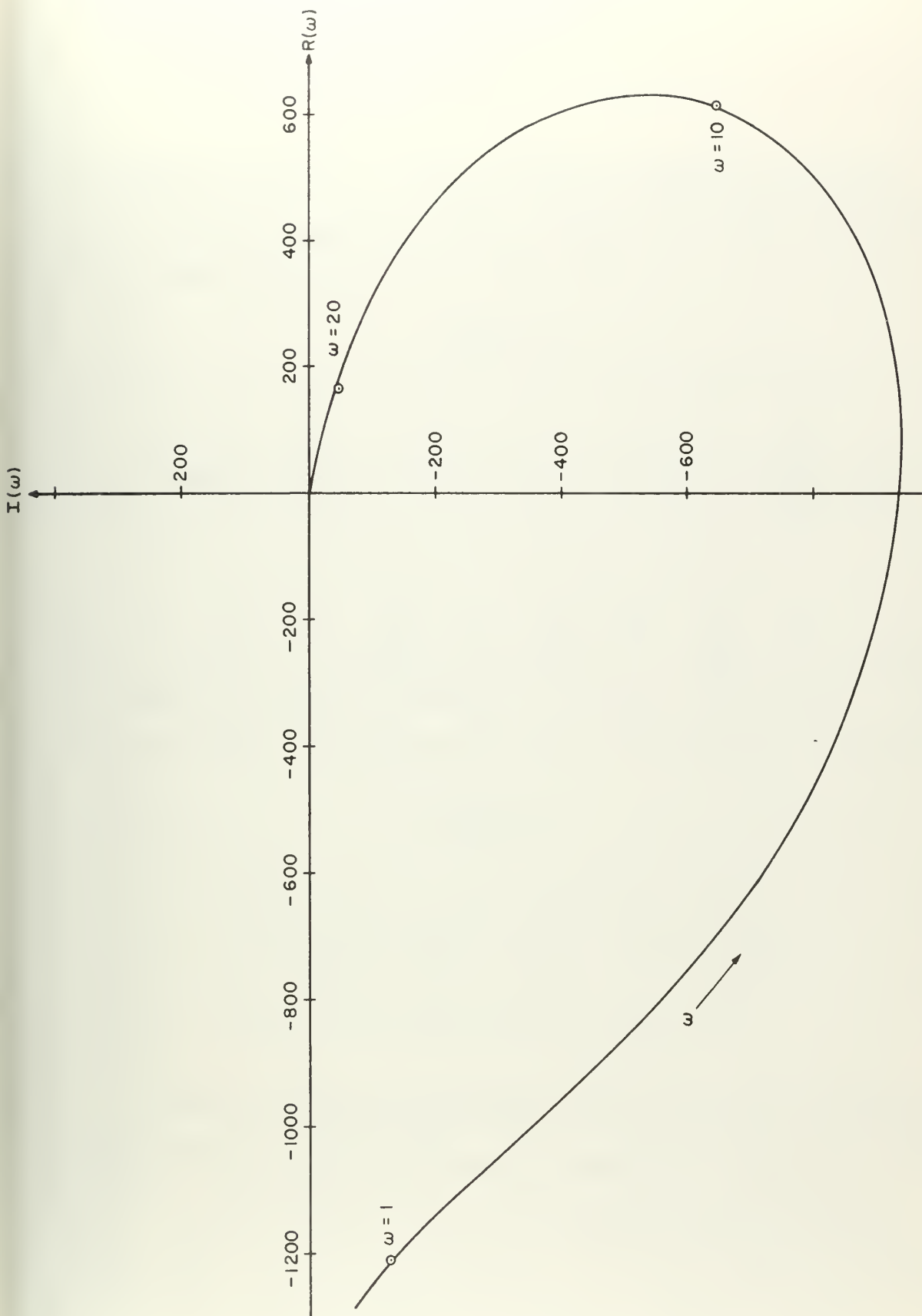


Fig. 5-7 : Open Loop Unstable Plant ($\zeta = -0.35$, $\omega_n = 7.2$)

for selected frequencies, in Tables 5-4 and 5-5. A glance at the row for $-L_n$ in these tables explains why Figs. 5-6 and 5-7 are presented qualitatively, i. e., the scale is such that the universal overlays, given with this thesis for the Ω regions, cannot be used.

TABLE 5-4

| ω | .5 | 1 | 2 | 3 | 4 | 5 |
|-----------------|------------|-----------|----------|-----------|------------|------------|
| $-L_n(j\omega)$ | -288-j83.6 | -246-j164 | -78-j243 | 43.8-j167 | 62.2-j89.2 | 54.8-j51.2 |

| ω | 10 | 15 | 20 | 30 |
|-----------------|-----------|----------|----------|---------|
| $-L_n(j\omega)$ | 17.6-j6.6 | 8.1-j2.0 | 4.61-j.8 | 2.1-j.2 |

TABLE 5-5

| ω | 1 | 10 | 20 |
|-----------------|----------------|--------------|-------------|
| $-L_n(j\omega)$ | -1214 -j122 | 617 -j647 | 165 -j48 |

In the face of an open-loop-unstable plant the plots must be interpreted with even greater care than before. The successful application of the variable parameter mapping technique to systems having open-loop instabilities may be of direct use in extending this study to the monitoring situation where the human operator is called upon to control a destabilized plant (or plant plus compensator). In the normal Nyquist plot, if a system is open-loop unstable, e. g., has two right-half-plane poles, then for the system to be closed-loop

stable the net number of encirclements of the critical point must be exactly equal to the number of those right-half-plane poles and the encirclements must be in a CCW sense. This results directly from the number of encirclements yielding the difference between the number of zeros and poles of the closed-loop-system-function denominator (characteristic equation when set equal to zero). Here, zeros (CW rotations) are regarded + and poles (CCW rotations) are regarded -. For no zeros to be present in this example, the net number of encirclements $N = Z - P$ must be $N = 0 - 2 = -2$. That is, for stability, the plot of the so-called "difference vector" as s traverses the entire right-half plane must undergo two complete counterclockwise revolutions about its origin (or, equally well, one CCW revolution for the positive frequency branch, since there is symmetry about the real axis). If $N = 0$, there are also two zeros in the right-half plane and the system is closed-loop unstable. What does a CCW encirclement appear as in the dual Nyquist plots? The single parameter variation as in Chapter Four and Figs. 4-8 through 4-12 may be viewed to gain some insight into this question. Specifically, suppose the situation were as depicted in Fig. 5-8, where there is shown a hypothetical plot of some $-L_n$. Clearly, if $\omega_{f1} > \omega_{v1}$, there is a CCW encirclement, while if $\omega_{f1} < \omega_{v1}$, there is no encirclement. If attention is now turned to Smith's results, it can be seen from the data points of both unstable plants that there are no encirclements, and since Smith used an unstable second order plant (two right-half-plane poles), it appears that this is an indication of two zeros in the right-half-plane and closed-loop instability. In helping to interpret the somewhat enormous plot of Fig. 5-7 ($\zeta = -0.7$), the following quote from (1) appears pertinent. "Controlled element gain had no apparent effect on the operator's ability to stabilize the system. Changes in K (K_p in this thesis) from 300 to 900 were barely noticeable by the operator at $\omega_n = 2.5$ rad/sec." For the $\zeta = -0.35$ case, the following comment appears. "The controlled element gain, K , (K_p in this thesis) was originally set at 200. The final value was 1200. Changes in gain produced no apparent change in the operator's technique. The only noticeable effect produced by gain changes was to decrease the maximum control deflection." Since the Smith study addressed itself to the control problem presenting extreme difficulty to the human operator, i. e.,

(1) Smith, R. H., An Experimental Determination of the Limits of Pilot Controllability for Unstable, Oscillatory Second Order Systems, Thesis for degree of Aeronautical Engineer, Univ. of Cincinnati, 1961 (also appeared as a paper, On the Limits of Manual Control, IEEE Trans. on Hum. Factors in Electronics Sept. 1963.)

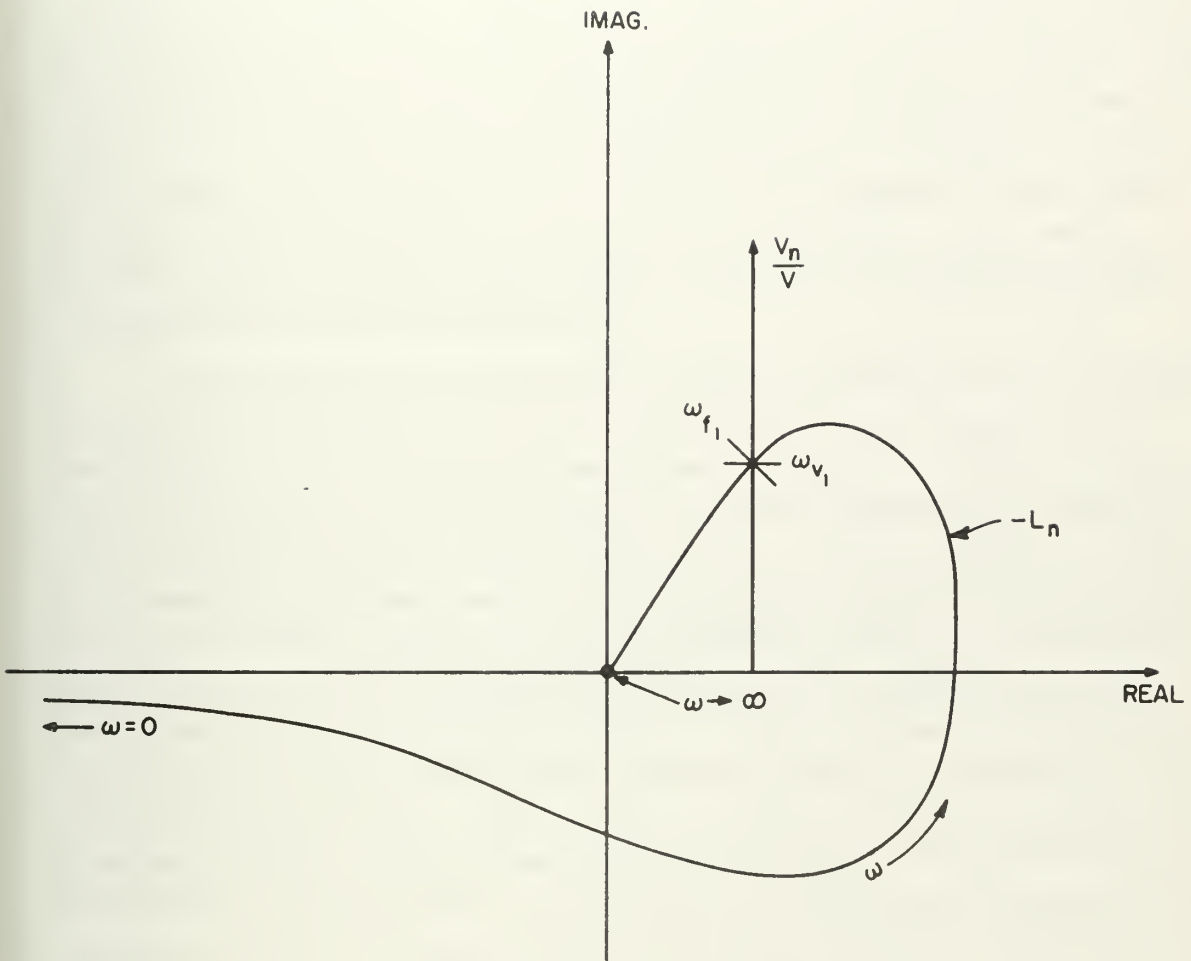


Fig. 5-8 : Interpreting Open-Loop Instability Effects

controlling an unstable-open-loop system, one further quote from (1) is appropriate. "Observations and operator comments during the conduction of the experiments for $\zeta = -0.35$ indicate the appearance of a very important phenomenon. In watching the operator at his control task, it seemed as though his reactions were intuitive. The changes in error were so fast and the error rates were so great that satisfactory control looked impossible; yet, somehow, sufficient stability was generated by the operator. After becoming accustomed to the control task the operator often remarked that he was simply outguessing the system. He made the statement that he was attempting to get his stick motion in phase with the error and trying to keep it there, i. e., system stability was his foremost requirement -- keeping error minimum was not consistent with maintaining stability."

The task of interpreting Figs. 4-6 and 5-7 (actually the data of Tables 5-4 and 5-5 in conjunction with Fig. 4-22), must now be considered; it is obvious that to stabilize the systems all the human gain must have been quite low. Unfortunately, the Smith study did not provide information relating to K_h since Smith's stability criterion was taken to be positive phase margin for a corresponding amplitude ratio in terms of K_h ; the calculation technique is not clear. The upshot, however, as recognized also by Smith, is that the generalized transfer function for the human operator (2-1) is a model incapable of stabilizing the systems under consideration. Yet the human operators in the experiment were able to stabilize the system (5-11) for the conditions (5-9) and (5-10), though just barely. That is, these results are supposed to reflect the "limits" of manual control. The plant characteristics (5-9) and (5-10) define the boundary of marginality, i. e., that point at which the human operator can no longer control the system, for a plant having form (5-11). Greater confidence accrues to the predictive capability of this theoretical approach if it too displays such marginal conditions. In order to account for the experimental results, Smith hypothesized a second lead term for the human operator model so that stability results from the theoretical description, also. This tack is

(1) Smith, R. H., An Experimental Determination of the Limits of Pilot Controllability for Unstable, Oscillatory Second Order Systems, Thesis for degree of Aeronautical Engineer, Univ. of Cincinnati, 1961 (also appeared as a paper, On the Limits of Manual Control, IEEE Trans. on Hum. Factors in Electronics, Sept. 1963.)

not taken here. It is sufficient to demonstrate that the techniques of this thesis imply a marginal stability situation. Thus, for $\omega = 10$ and $K_h = 0.1$ (not an infrequent value for human-operator gain) from Table 5-4, $-L_n = 1.8-j.7$, and it is conceivable that if K_h were slightly smaller than 0.1, the $-L_n$ locus might intercept the small semicircle (which is not a function of frequency, Table 4-3) projecting below the real axis on all the Ω plots. Under such circumstances, there is an enclosure and concomitant CCW encirclement indicating closed-loop-system stability. If it is felt that this discussion affords nothing more than an amateurish job of rationalization, it is also evident that the plots seem to indicate no clear cut conditions of one kind or the other. This substantiates that the mapping analysis adequately depicts the marginality of the situation. Application of the variable parameter mapping technique to the Smith data certainly demonstrates that the human operator is being pressed to the limit of his abilities.

The possibility of applying the techniques of this thesis to the multi-axis situation should not be discounted, i. e., the study provides a theoretical method of approaching the multi-axis situation where a particular "nominal" transfer function may not apply with invariance over all channels. The plots display the entire adaptive range of the human-operator transfer function and, since the bandwidth dealt with is so limited ($0 \leq \omega \leq 20$), the number of plots required to cover the necessary frequency range is pleasantly restricted. This may prove also, to be a more expeditious and fertile approach even for the single axis problem, since the automatic part of the system and the human operator are displayed separately. This isolation permits observation of the control characteristics of the automatic part of the system from the same plots used with the human controller analysis. Indeed, the positioning of the $-L_n$ locus relative to the $(+1+j0)$ point gives precisely the information normally obtained in a standard Nyquist analysis with respect to the $(-1+j0)$ point.

No experimental monitoring studies as such have been available to the author and none exist to his knowledge. Information vital to multi-axis monitoring by human operators may devolve from the Elkind studies of human response to abrupt plant changes, i. e., sudden changes in the gain and polarity

of a controlled element that was simply a gain; however, considerable experiment with more complex situations is needed to generate confidence in any inferences. ⁽¹⁾ ⁽²⁾ The theory presented herein works remarkably well for the single-axis situation and an equivalent simplified monitoring topology with open-loop instability, i. e., the theory predicts results consistent with the empirical evidence. Experimental results are needed to determine if the theory is totally applicable or, at the least, can continue to provide a valuable tool for design insight and "rule of thumb" criteria in multi-channel task situations.

The difficulty with the experimental studies reviewed here is that all parallel compensators, G_c , have been equivalently zero, Fig. 2-2. As indicated above there are no experimental results for monitoring. The question is: can theory predict that the human operator will have a less difficult task in a monitoring situation and how can this be determined? It would be instructive to observe an H_o controlling a G_p with $G_c = 0$, then compare the human effort there with that required when an H_o is paralleled by a G_c designed for good automatic control of the same G_p . Such information is not available at present. The problem is compounded by the valid question as to what role the human would play when the system is performing properly, i. e., presumably the human enters only upon some catastrophic failure of one kind or other. Obviously, considerably more research is required in this area.

5.5 Summary

The variable parameter mapping of the three time constants (which man can adapt to a given task) in the human-operator description, is used in conjunction with a slightly modified Nyquist plot of the plant characteristics to

(1) Young, Green, Elkind and Kelly, Adaptive Dynamic Response Characteristics of the Human Operator in Simple Manual Control, IEEE Trans. on Hum. Factors in Electronics, Sept. 1964.

(2) Dander V. A., Predicting Pilot Ratings of Multi-Axis Control Tasks from Single-Axis Data, IEEE Trans. on Hum. Factors in Electronics, Sept. 1963.

demonstrate that stability of the closed-loop system (with a man in the loop) can be predicted. Experimental verification of the theory is obtained by analyzing certain systems which have been studied empirically in the past and for which results have been published in the literature. To this end, transparent overlays are provided which facilitate stability determination. The theoretical predictions parallel the experimental results remarkably well; one of the sets of data relates to a system which is open-loop unstable. The correlation possible between boundary lines of the mapping and the physical capacity of a human operator is cited and, in the cases of stable control, the margin of human capacity remaining (displayed on the same diagram) is indicated.

CHAPTER SIX

PERFORMANCE6.1 Introduction

As in the discipline of inanimate systems engineering, several performance measures have been used for man-machine systems to provide a so-called figure of merit or index of quality and, as in the case of the all-hardware system, deficiencies have been found to exist with all these scalar indices. The difficulty in attempting to quantify a complex system's performance by virtue of a single number is that satisfying the requirement on this unique scalar may still result in a system with other unsatisfactory characteristics. Of course, the designer of performance measures answers that all pertinent characteristics are duly weighted and included in the scalar index thus causing it to reflect those properties deemed important. Still, the final selection of those properties regarded as necessary for an acceptable system and assignment of relative weights are only two of many reasons for viewing single performance measures with some caution. Generally, the selection of performance index has often been predicated more on ease of calculation and other conveniences than upon how accurately the truly efficacious characteristics of a system are reflected. In the psychological literature, time on target, an easily obtained number has been used extensively although studies have shown its deficiencies; it has even been demonstrated that there is considerable danger in inferring that system-measure changes imply system-component changes.⁽¹⁾⁽²⁾ Other indices used by psychologists studying manual control have been average error, average-absolute error and root-mean-square error; attempts

(1) Bahrick, H. P., P. M. Fitts, and G. E. Briggs, Learning Curves-Facts or Artifacts, Psych. Bull., 54, 256-268, 1957.

(2) Taylor, F. V., and H. P. Birmingham, That Confounded System Performance Measure-A Demonstration, Psychological Review, 66, 178-182, 1959.

have been made to measure task difficulty but independent critics have dismissed them as measuring little about the specific requirements of a task.⁽¹⁾ The engineering approach to the performance measure quandary has been to employ the seemingly dependable (and quite convenient) mean-square error, as well as certain other control theory measures. The period of time over which the measures themselves are calculated serves to influence the result and thereby confuse the problem still further. Historically, pilot opinion has been used as an acceptance technique in the case of aircraft flight control systems. In this regard, the following quote is considered appropriate. "It should be clear at this point that a single measure for a complex tracking situation cannot give information concerning both system performance and operator performance" (underlining theirs).⁽²⁾

In this chapter, the various performance measures are discussed and a brief survey is given of some techniques available for "optimizing" a system so as to best satisfy these indices. The formulation of a "vector" performance index is considered as a means of circumventing certain of the disadvantages of scalar indices. The concept of a "best-overall" compensator (fixed compensation), for a plant with variable dynamics, is introduced and a technique for specifying the parameters of this compensator is offered. This technique depends upon the formation of an "artificial" function which, when minimized to zero, identically satisfies a set of performance requirements. A "direct search" procedure, used in conjunction with this "artificial" function, which is sequentially modified, serves to isolate satisfactory values for the compensator time constants. In this manner, the space of acceptable compensator time constants is defined and correlated with the space of required performance. The entire procedure is especially amenable to the electronic computer. An example of the search technique is given at the end of the chapter.

-
- (1) Taylor, F. V., Human Engineering and Psychology, in "Psychology: A Study of a Science", Koch, S., (Ed.), Vol. 5, McGraw-Hill, 1963.
- (2) Summers, L. G., and K. Ziedman, A Study of Manual Control Methodology with Annotated Bibliography, NASA Contractor Rpt. CR-125, Nov., 1964.

6.2 Varying Parameter Plants and "Best" Performance

The question frequently arises how "best" to use the human operator in any given system; it is assumed that the stability criterion is satisfied. Given the requirement for a human operator one may ask, alternatively, how to design a satisfactory compensator for use in a system which includes this human controller? The design task may be complicated further by specifying a plant which itself contains some varying parameter. Generally, the controls engineer has tended to avoid computer design for the more familiar Bode' diagram, frequency plot or root locus technique. An approach to the performance problem, is pursued in this thesis which, it is hoped, can be suggestive of a standard technique, or algorithm, for handling far more severely complex situations than can be treated here. The method is particularly applicable to systems containing human controllers since the human adapts his characteristics to some unknown internally-generated criterion (which is frequently taken to be some function, or functional, of the error). The situation is quite fluid in the face of systems with varying-parameter plants since the error itself then changes due to this parameter variation thereby causing further adaptation on the part of the human. It is felt intuitively that some design compromise on an adequate compensator should be possible so as to provide the "best" performance over some range of parameter variation. The question immediately arises regarding a proper criterion (performance index); this is a subject around which much controversy still rages within the engineering profession.⁽¹⁾⁽²⁾ Frequently, the physics of the situation suggests a performance index which leads to a mathematically intractable problem. The designer then must compromise on a choice of performance index which permits a meaningful solution, i. e., does not preclude analysis, yet still provides a useful quality mea-

(1) Newton, G. C., L. A. Gould, and J. F. Kaiser, "Analytical Design of Linear Feedback Controls", John Wiley & Sons, New York, 1957.

(2) Gibson, J. E., "Nonlinear Automatic Control", McGraw-Hill, 1963.

sure for the system. Usually, the indices yielding the greatest analytical ease in their calculation are the mean-square error and the integral-square error. For inputs having Gaussian characteristics and certain other properties which are frequently encountered (or at least frequently assumed) it can be shown that several error criteria offered as more appropriate lead to the same results as the mean square error criterion.⁽¹⁾ Of course, the advent of high speed computation and concomitant expeditious numerical integration offers great latitude in the choice of a performance index.

Some work has been done indicating that minimization of the mean-square error is a fair approximation to what the human controller attempts, i. e., it appears the ultimate effects are similar.⁽²⁾ Generally, the mean-square error is used as a criterion for stochastic inputs; indeed, this is the basis upon which the human-operator transfer function has been generated.⁽³⁾ For transient inputs, the analyst frequently uses the integral square error criterion; of course, the ISE is used only in the design of systems having zero final error since continual finite error leads to the non-convergence of the ISE index. With either measure, however, stability checks must be made at discreet (also discrete) moments to insure that the values of the adjustable parameters which minimize the performance index correspond to a stable solution. This is elaborated upon below. Thus, one finds that the MSE is used as an index of, loosely speaking, steady state performance while the ISE is used to measure transient performance characteristics. Certainly, the selection of a performance index is at the discretion of the designer (or the customer) and modern digital computation offers the possibility of numerically computing heretofore mathematically unattractive but physically useful response criteria.

(1) Sherman, S., Non-Mean-Square Error Criteria, IRE Trans. on Information Theory, Sept., 1958.

(2) Roig, R. W., A Comparison Between Human Operator and Optimum Linear Controller RMS-Error Performance, IRE Transactions on Human Factors in Electronics, Mar., 1962.

(3) Graham, D., and D. McRuer, "Analysis of Nonlinear Control Systems", John Wiley & Sons, 1961.

The remainder of this section is confined to the exposition of a possible approach to solving such a "best" performance problem. The attack is along conventional lines while that of the next section is considerably more original and exotic. The system studied here is somewhat simplified but it should be recognized that the simplification incurs no loss of generality.

The system under investigation appears in Fig. 6-1. The configuration within the dotted lines represents a "choice" situation, i. e., the human operator either chooses to operate upon the error with his dynamics or he does not. Again, such a situation where the human operator decides whether or not to enter the circuit may be considered as an alternate form of the monitoring situation. The representation here does not place the human operator in parallel with the original compensation but in cascade. It is recognized that the design of G_c for good automatic control may be distinct from a good design with a human controller in the system; therefore, the human is cascaded always with some G'_c such that $G_{cT} = G'_c G_c$ provides good compensation for the human controller-plant combination. It will be observed that no loss of generality occurs if the situation is depicted as in Fig. 6-2 since the plant-human path is under study. The human-operator transfer function is considered in the simplified form

$$H_o = \frac{e^{-.1s}}{s+h} \quad (6-1)$$

where h can be varied adaptively over some range. The unknown compensation to be determined is cast in the form of a straight gain, α , and the plant has the form

$$G_p = \frac{K}{s^2 + \beta s + 1} \quad (6-2)$$

where β varies over some range. It should be noted as the study continues that any class of more complicated frequency-sensitive compensation, a plant with several varying parameters or a mathematical model for the human operator containing more adaptive variables, only makes the calculations more severe but does not change the general technique. Thus, Fig. 6-2 takes on the form of Fig. 6-3.

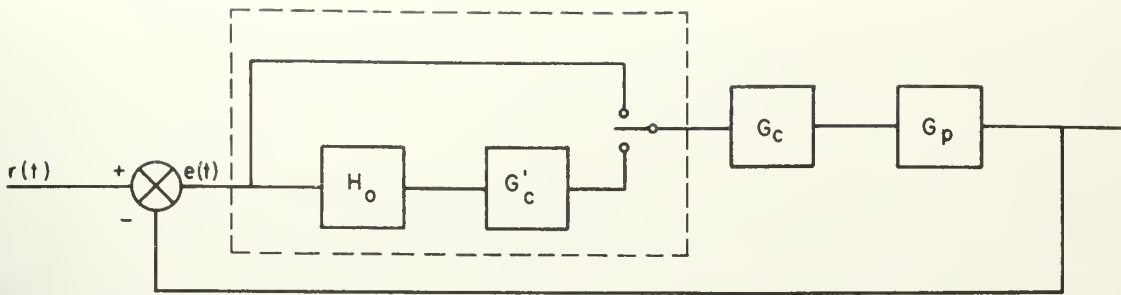


Fig. 6-1

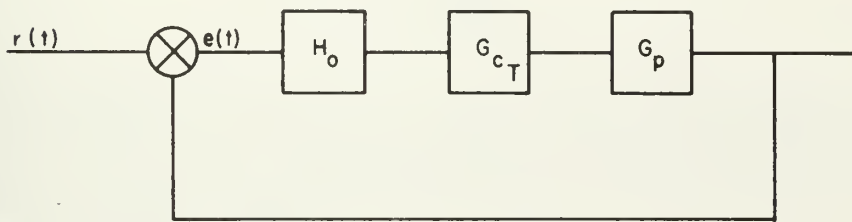


Fig. 6-2

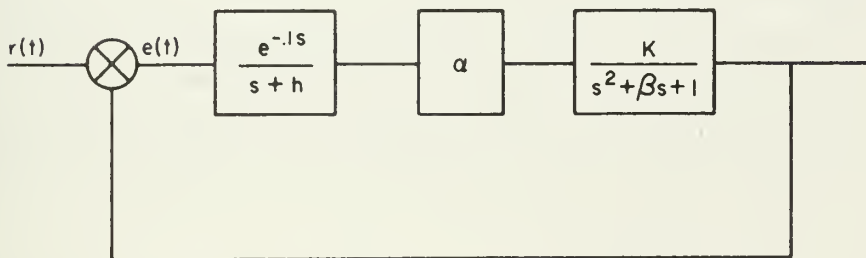


Fig. 6-3 : Human Control of Varying-Parameter Plant

Since the study is restricted to deterministic inputs (steps, ramps, etc.) the performance index is chosen to be the ISE,

$$\text{ISE} = I_e = \int_0^{\infty} e^2(t) dt \quad (6-3)$$

where $e(t)$ is the error signal.

A design is desired, i. e., a value for α , such that the ISE is minimized. The standard analytic technique for solving this problem uses a rational transform approximation for $e^{-\tau s}$

$$e^{-\tau s} \approx \frac{1 - \frac{\tau s}{2}}{1 + \frac{\tau s}{2}} \quad (6-4)$$

This permits evaluation^{*} of the ISE as a function of the free parameters. One may then attempt to minimize the integral by taking derivatives of ISE with respect to each of the free parameters setting them equal to zero and solving for the critical value of each parameter. Thus

$$\frac{\partial I_e}{\partial h} = 0, \quad \frac{\partial I_e}{\partial \alpha} = 0, \quad \frac{\partial I_e}{\partial \beta} = 0 \quad (6-5)$$

The step above is often more easily described than accomplished, but the main difficulty is not merely that of explicit solution. In the case with a human in the system, one must assume there is an attempt to adapt operating procedure so as to minimize some function of the system error, irrespective of the value of α and β . The β variation is not controlled and may represent many possible

* The evaluation is easily done by means of tables. (1)

(1) Newton, G. C., L. A. Gould and J. F. Kaiser, "Analytical Design of Linear Feedback Controls". John Wiley & Sons, New York, 1957.

phenomena; e. g., change of response characteristics with altitude by a high-performance aircraft, drift in plant parameters, etc. The problem may be approached as one wherein Nature is in direct opposition to satisfactory performance and a design is sought such that performance is at least as good or better than some value of the index of merit, in this case, ISE. For tutorial purposes, consider α and β , each having such range that they are adequately described by a grid of five (5) values, each grid covering the respective ranges of α and β . Thus, for each pair (α, β) there is obtained that value of h^0 which minimizes the general performance index S ; this index is a function of the three variables α, β, h and is taken in ISE form. The minimization of

$$S(\alpha, \beta, h) = \int_0^{\infty} e^2 dt \quad (6-6)$$

is then given notationally by

$$1. \min_h \int_0^{\infty} e^2 dt \xrightarrow{\text{produces}} h^0(\alpha, \beta) \quad (6-7)$$

i. e., for each (α, β) over the chosen grid

$$\begin{array}{lll} (\alpha, \beta) = (0, 1) & \xrightarrow{\text{one finds an}} & h^0_{01} \\ & \longrightarrow & h^0_{02} \\ & \vdots & \\ & \longrightarrow & h^0_{ij} \end{array} \left. \vphantom{\begin{array}{l} h^0_{01} \\ h^0_{02} \\ h^0_{ij} \end{array}} \right\} 25 \text{ } h^0\text{'s}$$

This step represents the effect of assuming that the human operator always does the "best" he can, where "best" is the optimization of some error criterion (here the identical performance index, ISE, is retained throughout as the criterion but this is not mandatory, i. e., distinct performance indices can be minimized for h, α and β , respectively).

In the next step, for each value α can take on there is found that value of β which maximizes the ISE. Here one observes the result of assuming the "cussedness" of the plant; this is, at times, called "worst case design". The disadvantage of this approach lies obviously with the extreme conservatism of the final solution. Thus,

$$2. \max_{\beta} \int_0^{\infty} e^2 dt \longrightarrow \beta^0(\alpha) \quad (6-8)$$

i. e., for $\alpha = 1$, run through $\beta = 1, \dots, 5$ to get an optimal β_1^0
 \vdots
 $\alpha = 5$, run through $\beta = 1, \dots, 5$ to get an optimal β_5^0 } $5 \beta^0$'s
 where the "optimal" β maximizes $\int_0^\infty e^2 dt$.

Finally, one looks at the five values of (6-6) for $\alpha = 1, \dots, 5$ and picks the minimum of these, thus yielding the "best" α . Hence,

$$3. \min_{\alpha} \int_0^\infty e^2 dt \longrightarrow \alpha^0 \left\{ 1 \alpha^0 \right. \quad (6-9)$$

The above procedure insures that performance

$$S(\alpha^0, \beta, h^0) \leq \text{value} \int_0^\infty e^2 dt = S(\alpha^0, \beta^0, h^0) \quad (6-10)$$

where β^0 , it will be recalled, refers to the "worst" situation. For those familiar with Game Theory, the use of a game theoretic approach in the technique is noticeable, i. e., a minimax solution is being obtained. Clearly, even in this extremely simplified example many calculations and evaluations are required and frequent comparisons between them must be made. The technique is especially suited for digital or other machine computation.

Two additional points with regard to this approach should be mentioned. The first step (the adaptive action) yields an $h^0(\alpha, \beta)$ from which one can proceed to evaluate $S(\alpha, \beta)$ as indicated above, but care must be taken to insure that $(\alpha, \beta) \in R_s$; that is, the pair (α, β) must correspond to a region of stability. Otherwise, $S(\alpha, \beta) \rightarrow \infty$ if $(\alpha, \beta) \notin R_s$. The parameter space plot of Fig. 6-4 should clarify this. If the region of stability has the form sketched in $\alpha\beta$ space and the range on α and β is indicated by lines (1)-(2) and (3)-(4), respectively, then it is clear that during the step $\max_{\beta} S(\alpha, \beta)$ for the value $\alpha = \alpha_u$ the criterion is maximized by picking $\beta = \beta_u$, since this combination corresponds to an unstable system. Either a manual check must be conducted to ensure that the optimal values determined correspond to a stable condition or the computer program must

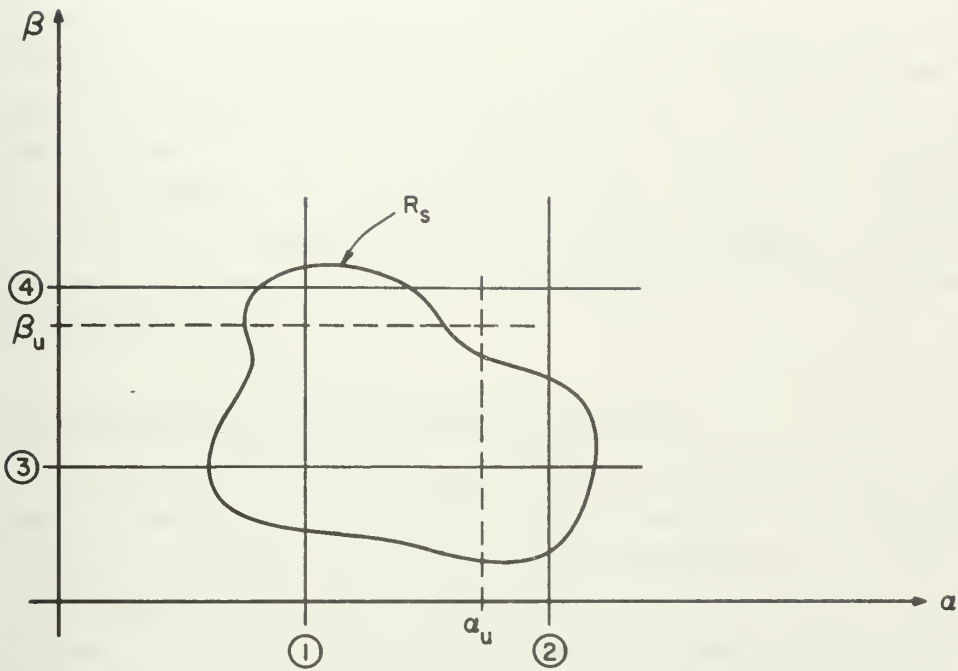


Fig. 6-4 : Solutions Outside the Region of Stability

include a check for stability during the computational cycle.

If the transform of $e(t)$ is a ratio of polynomials then the evaluation of ISE is straightforward, by virtue of previously prepared tables, and only the minimization is carried out by computer. If $e(t)$ transforms into an irrational form (as it does with a pure time delay) then evaluating the ISE is not so simple. The integral may be truncated, however, and evaluation carried out numerically. To follow such a procedure, however, reduces any feeling of commitment to ISE, or other mathematically attractive figures of merit, since their calculation would require as much work as that involved in other more physically appealing performance indices. Also, a coordinated computer program may be possible such that, while the performance index itself is being calculated, the information generated in that process is used to simultaneously minimize the index and provide a stability check.

6.3 The Concept of a "Vector" Performance Index

A technique is now considered for determining the "best-overall" compensator for any given system with a human controller. The approach in this section may be used for a plant with variable parameters; the only restriction is the usual aggravation imposed by the "curse of dimensionality", i. e., the larger the number of varying parameters the more iterations required.⁽¹⁾ The solution to be presented depends intimately upon an initial stability study conducted in accordance with the procedure previously outlined in this thesis.* Thus, for the situation of Fig. 6-5, the result desired is values for K_c , α and β which provide good, overall performance, defined by specifications on the range of the

(1) Bellman, R., "Adaptive Control Processes-A Guided Tour", Princeton Univ. Press, Princeton, New Jersey, 1961.

* The author has been unable to find in the literature information relating to the step response of a human-operator-frequency-sensitive-plant combination, although there is considerable published on such responses for plants represented by pure gains. Thus, the study represents a theory which is yet to be checked empirically, although all previous experiments would seem to justify this theoretical extension.

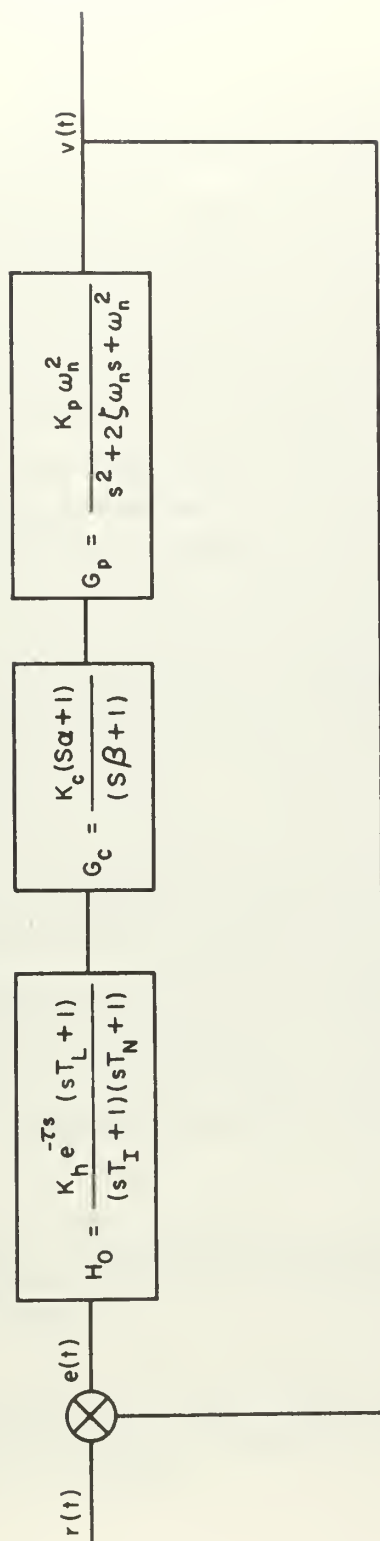


Fig. 6-5 : Human Control with "Best" Compensation

triple (t_r, o_s, t_s) , in the face of a plant which has a varying ζ . The triple (t_r, o_s, t_s) represents the values of rise time, overshoot and settling time, respectively, which follow in accordance with some arbitrary definition. Of course, the triple may be regarded as a vector quantity in three space^{*}. The stability analysis outlined in Chapters Four and Five is first accomplished; for the class of compensator being considered, one zero and one pole, it can be determined quickly what approximate ranges of values on α and β provide good stability. The value of K_c is not difficult to determine since a nominal value of K_h may be used, thus indicating quickly what the overall loop gain should be, from this a first approximation to K_c may be calculated. Refinement of this value is made along with the search for final values of α and β . Gain adaptation is the easiest act for the human operator to perform; indeed, his task is regarded as simplest when he acts merely as an amplifier.⁽¹⁾ This aspect of the variation therefore can be discounted in the remaining analysis. The competent designer desires a final empirical test anyway, if only by simulator, overall operation can be double checked at that time. In Section 6.5 an example is given which includes consideration of K_c .

In addition to satisfying the general stability requirement the designer does, of course, attempt to use the freedom available in selecting α and β so that the plant's Nyquist curve passes through that part of the Ω region corresponding to reduced task requirements on the human operator, (Chapter Five). The entire analysis so far has been conducted for a nominal value in the range of variable ζ . Generally, there are some ranges on α and β which seem to meet the conditions, or it is apparent they are impossible to meet under the constraints imposed, or a

* Quality measures other than the three suggested are, of course, possible; the technique permits considerable latitude in employing multidimensional performance measures.

(1) Birmingham, H.P. and F. V. Taylor, A Human Engineering Approach to the Design of Man-Operated Continuous Control Systems, NRL Rep. 5333, 7 Apr., 1954 (also published as A Design Philosophy for Man-Machine Control Systems, Proc. IRE, Dec., 1954.)

different class of compensator may be required. With the ranges of α and β that provide acceptable stability available, nominal values for the human controller can be read from the plot using the time-invariance assumption. These values must also retain the "high-frequency" end of the compensated Nyquist contour in a location satisfying the stability criterion, as reflected by the corresponding Ω region. * That is, the triple (T_L, T_I, T_N) must have value such that it lies to the left of the plant's Nyquist curve and still be within the range of human adaptivity. Simultaneously, T_L, T_I and T_N should be as small as possible; particularly T_L , since this parameter is almost a direct indication of "comfortable" adaptation. All these conditions must also be consistent with physiological realities.

With the analysis above conducted for a nominal ζ , it is then repeated for a grid of ζ values covering the range of ζ . The intersection ** of the various ranges on α and β yields those ranges of α and β which satisfy stability over the expected range of variation in ζ . These final ranges of acceptable α and β must now provide the specific values of α and β which yield the best performance. Again, a vector (rather than a scalar) definition of performance is being used. The requirement is to meet specifications on rise time, t_r , overshoot, o_s , and settling time, t_s . There is no problem in determining what values of t_r, o_s and t_s result from any given set (α, β) ; the question is how to go about selecting subsequent trials of α and β so that each iteration is at least as good as the last. What is sought is some "scientific method" for approaching the sub-space of acceptable performance within the space of acceptable stability. The following approach provides a routine technique, or algorithm which lends itself to computer solution. With the resolution of ranges on α and β which provide satisfactory stability characteristics, there is conducted a search for the sub-range

* Inherent in this statement is the assumption that this is what the human does; i. e., it is hoped this adequately reflects his unknown performance criterion.

** That range over α and β , respectively, whose end points are within all the other ranges of α and β determined for various ζ is called the intersection (greatest common range of α and β over ζ). It should be noted that this step considerably reduces the experimental region remaining.

which yields good performance, Fig. 6-6(a) and (b). Quite possibly none may exist; this situation is discussed later. A grid search of $\alpha\beta$ space could be performed to find acceptable (t_r, o_s, t_s) , but that is a strictly trial and error method. If the (α, β) range is small, straightforward grid search could be the most efficient technique; if there are more than two parameters available for manipulation, however, the algorithm is probably more efficient.

The function

$$f = (t_r - T_r)^2 + (o_s - O_s)^2 + (t_s - T_s)^2 \quad (6-11)$$

is formed where $0 < t_r \leq T_r$, $0 < o_s \leq O_s$, $0 < t_s \leq T_s$, i.e., T_r , O_s , and T_s are the maximum acceptable bounds on t_r , o_s and t_s , respectively. It should be noted that the function f has the interesting property that when it is zero the specifications are met identically at their upper bounds, i.e., α and β have values such that the surface boundary of the acceptable performance region in (t_r, o_s, t_s) space has been attained, Fig. 6-6(b). The function is an artifice created specifically to facilitate finding some point on this surface and, in this instance, is still a function of two independent variables α and β , although its ultimate evaluation requires three numbers t_r, o_s, t_s . Thus, it is desired to minimize f over the space of $\alpha\beta$ which yields stability, recognizing that due to the sum of squares form of f it cannot be negative, while to meet specifications just barely, $\min f$ must equal zero. The form of the function therefore precludes the occurrence of saddle points when the min of zero is attained. The search techniques for "optimum" α and β must incorporate certain characteristics to keep the designer adequately informed.

First, it is desired to check each of the terms $t_r - T_r, o_s - O_s, t_s - T_s$ algebraically; in the event that they are all negative the program should STOP and PRINT OUT such information since this is a reflection of an adequate design whose response falls inside the acceptable performance region. Second, when the search pattern finds $\min f$ a PRINT OUT should again be required, since it must then be determined by the designer whether the results are sufficiently close to zero to warrant an interior search of the acceptable performance region. If they are not, this is an indication that the requirements on t_r, o_s, t_s cannot be

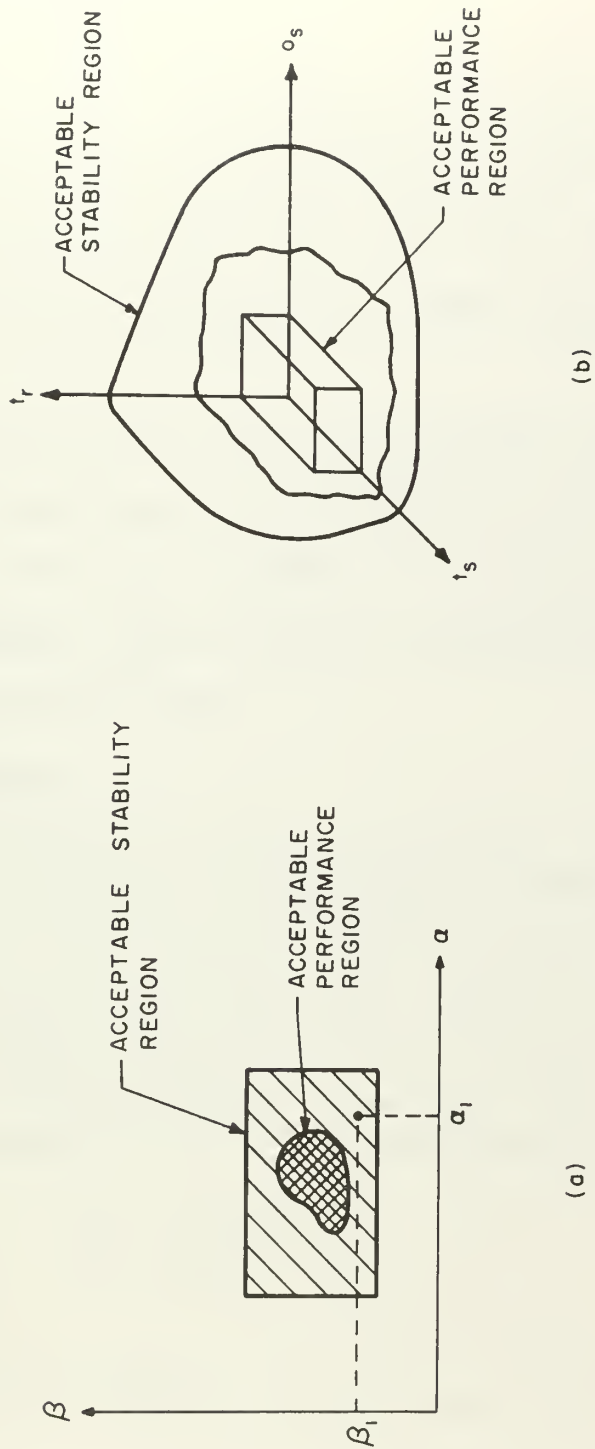


Fig. 6-6 : Stability and Performance Regions in Transform Domain (a) and Time Domain (b).

met and a different class of compensator may have to be used or other compromise fashioned. The utility of such information should not be regarded lightly since it represents an analytical demonstration of inconsistent or impossible specifications. Too frequently there have been attempts at building systems to meet such impossible requirements and only successive futile results have led to their abandonment; any technique which uncovers such sets of impractical demands serves the important purpose of precluding the expenditure of funds and effort on an unrealizable quest.

Having located the surface of the acceptable region of performance, one proceeds in the "same direction" until the opposite side of the region is found in which case the entire line between (and the associated values of α and β) must be contained inside. Many times the "opposite side" is known due to constraints on the parameters, e.g., α and β must be ≥ 0 . In this manner, the region can be crudely, or completely, defined by successive computation along lines perpendicular to the one originally found.

6.4 The Gradient Method

Frequently, a steepest ascent (descent) technique is employed in such problems.⁽¹⁾ This type of approach is described briefly, although it will not be used; the description allows comparison with the procedure to be given in detail somewhat later. From (6-11) for an arbitrary (α_1, β_1) the function $f(\alpha_1, \beta_1)$ is determined, i.e., for the values (α_1, β_1) the system response to the input function of interest (here a unit step) is calculated and t_r, o_s, t_s is found thereby permitting f to be enumerated. For some $(\Delta\alpha, \Delta\beta)$ the functions $f(\alpha_1 + \Delta\alpha, \beta_1)$ and $f(\alpha_1, \beta_1 + \Delta\beta)$ are determined similarly. The function below is formed

$$\frac{f(\alpha_1 + \Delta\alpha, \beta_1) - f(\alpha_1, \beta_1)}{\Delta\alpha}$$

(1) Wilde, D. J., "Optimum Seeking Methods", Prentice-Hall, New Jersey, 1964.

and also,

$$\frac{f(\alpha_1, \beta_1 + \Delta\beta) - f(\alpha_1, \beta_1)}{\Delta\beta}$$

In this way, estimates are obtained on $\frac{\partial f}{\partial \alpha}$ and $\frac{\partial f}{\partial \beta}$ respectively.

But the gradient of f is

$$\nabla f = \frac{\partial f}{\partial \alpha} \bar{i}_1 + \frac{\partial f}{\partial \beta} \bar{i}_2$$

where \bar{i}_1 and \bar{i}_2 are unit vectors in the positive α and β directions, respectively. The gradient of f , of course, points in the direction of greatest increase of the function f . Since it is desired to minimize f the program proceeds in the opposite direction. By surrounding the point (α_1, β_1) by a pattern of, say, two points oriented as in Fig. 6-7, the gradient at each of these two points may then be determined and iteration continue in an appropriate direction. Thus, the direction of incremental change in α and β yielding the greatest decrease in f (or, increase in minus f) would be followed to obtain the next trial point. The search procedure is then repeated as one follows one's nose to a minimum. When the optimal point (be it min or max) is apparently reached it is good practice to conduct a search in the local area surrounding this isolated point to ensure that it is not an apparent optimum. For this purpose, non-linear exploration of the adjacent region is in order; Box and Wilson have discussed the problems inherent in such studies.⁽¹⁾

6.5 Direct Search and Pattern Strategy.

The upshot of the previous discussion is that any of several different type search techniques can be used to determine the "best" α and β . The approach used in this section, and which seems to offer the greatest advantage for the systems under consideration, follows that of Hooke and Jeeves.⁽²⁾ This technique of

(1) Box, G. E. P. and K. B. Wilson, The Experimental Attainment of Optimal Conditions, J. Roy. Stat. Soc., B13, 1(1951).

(2) Hooke, R., and T. A. Jeeves, Direct Search Solution of Numerical and Statistical Problems, J. Assoc. Comp. Mach.; 8, 2, Apr., 1961.

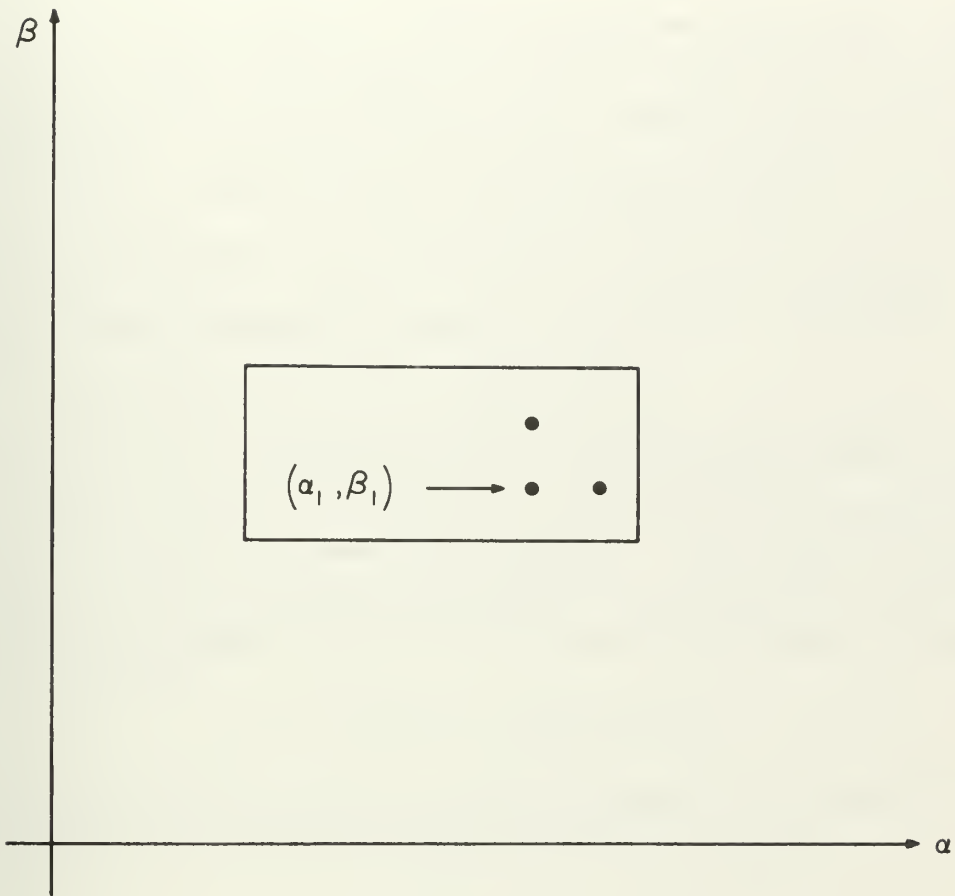


Fig. 6-7 : Perturbation Pattern for Gradient Determination

direct search with pattern strategy is especially suitable to minimization of a sum of squares, but its overriding virtues are its straightforward simplicity and compatibility with the electronic computer.

The starting techniques in pattern search are quite similar to those of the gradient or steepest ascent method, however, it appears at first glance that pattern search neglects some information that is readily determined and used in the gradient method. The effects of this neglect are more apparent than real. By "direct search" is meant a step-by-step examination of trial solutions and a method of comparing each to determine which is the "best" in some sense. Additionally, there is some method offered for choosing the next trial point. This last aspect forms the strategy part of the search procedure; a "pattern" approach is used and from this is derived the strategy's name.

The example to be used is shown in Fig. 6-8. It is assumed that the ranges of α and β have been discovered which yield good stability; specific values within that range are now desired which meet the step-response requirements on rise time, overshoot and settling time. The compensator gain K_c , is included as a third parameter whose value must be determined. The ranges K_c , α and β form a space of points representing possible solutions. Again the function, f , is formed

$$f = (t_r - T_r)^2 + (o_s - O_s)^2 + (t_s - T_s)^2 \quad (6-11)$$

For economy of notation, let the compensator gain, K_c , be denoted, K , then the point (K_2, α_2, β_2) is said to be "better" than (K_1, α_1, β_1) if*

$$f(K_2, \alpha_2, \beta_2) < f(K_1, \alpha_1, \beta_1) \quad (6-12)$$

The values K_2, α_2, β_2 are determined by adding some increment δ to each value K_1, α_1 , and β_1 ; the increment size used for each parameter need not be the same. This portion of the search procedure is called the exploratory phase. It is understood that by $f(K_1, \alpha_1, \beta_1)$ is meant the value of (6-11) found from the substitution

* The inequality is given for a minimization problem.

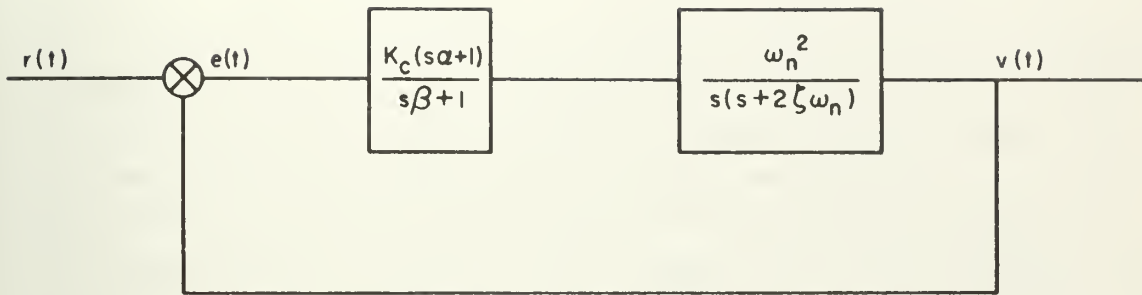


Fig. 6-8 : Compensation of a Varying-Parameter System

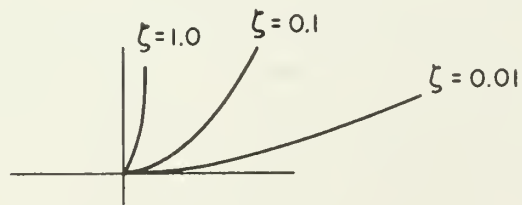


Fig. 6-9 : Sketch of $-L_n$ Plots

of the t_r , α_s and t_s associated with the system step-response when the compensator has gain and time constants K_i , α_i and β_i , respectively. The initial base point (K_1, α_1, β_1) is chosen arbitrarily or the investigator may exercise his "hunches". There may be none, one, or more points such that a min f of zero is found. The ramifications of this have been discussed above. Under the assumption (6-12) holds, there seems little reason in the face of such success, to alter the "direction of search"; therefore, exploration is continued by perturbing the independent variables with increments whose proportions are based on the number of successes achieved along each dimension in the recent past. This strategy forms the so-called "pattern" aspect of the search procedure; of course, great strides are made with pattern moves in relation to progress resulting from exploratory moves. An example is offered below. In those situations where a min $f=0$ is obtained, the direct search method continues to offer a consistent philosophy. The designer unquestionably desires additional information with regard to the size of the acceptable performance region and does not wish to stop at the barely acceptable surface bound, e. g., other satisfactory triples (K_i, α_i, β_i) may offer simpler design realizability. Direct search can forge across this region until it pierces the opposite surface; in this manner the performance region may be completely explored.

Termination procedures are required when a pattern move failure is followed by consistent failure from all exploratory moves. A retrenching is occasioned by reducing the nominal step size of the increments and new explorations are conducted; this is continued until a new pattern is constructed or the step size falls below some pre-selected margin. The lower bound on this margin is usually imposed by the demands of the particular computation interval peculiar to the computer used. Termination does not imply an optimal solution so the wary investigator tests the solution generated in radial directions about the final base point distinct from the coordinate axes since pattern search restricts itself to motion along these axes.

6.6 "Best-Overall" Compensation of a Varying Parameter System - An Example

The system for which a best-overall compensator is to be designed is

shown in Fig. 6-8. It should be noted that the human-operator transfer function has been omitted and a simplified plant assumed; the simplification is imposed for clarity only. In this manner, the techniques to be demonstrated are not obscured by technical complexities; no loss of generality occurs. The plant's damping ratio, ζ , is assumed to vary over the range $0.01 \leq \zeta \leq 1.0$; the rise time, t_r , is defined as the time required for the response to go from 10 to 90 percent of its final value; the overshoot, o_s , is the maximum departure from final value expressed as a percentage; the settling time, t_s , is defined as that time after which the response is confined to a tolerance band of ± 5 percent about its final value. In order to demonstrate the search technique for more than the two parameters α and β , the value of gain, K_c , is included as an unknown in the search.

Sketches of the negative Nyquist contours for the plant, $-L_n$, for values of $\zeta = 0.01, 0.1$ and 1.0 are shown in Fig. 6-9. The limiting ranges on α and β are taken to be $0 \leq \alpha \leq 5.3$, $0 \leq \beta \leq 25$. The variable-parameter mappings of Chapter Four can be used to aid in the design, since they are not required to represent the capabilities of a human operator, i. e., H_o has been assumed equal to unity in this example in order to illustrate the search method. Thus, the degenerate mapping shown in Fig. C-2, Appendix C, (for $\omega = 1.0$) can be used to represent the available region of variation for the compensator's time constants, α and β . Similarly, the corresponding regions on the transparent overlays (in the pocket of the back cover) can be used to provide additional frequency effects other than for $\omega = 1.0$.

It is evident from the overlays, used in conjunction with the more accurate plots in Fig. 6-10 (or root locus considerations), that lag network configurations ($\alpha < 3$) severely aggravate stability when ζ is very small, i. e., between $0.01 \leq \zeta \leq 0.1$. Appendix C contains further stability interpretation elaboration. From the overlays it is apparent that after introduction of a compensator, stability is possible for small ζ ($\zeta \approx 0.01$) only if the compensation is lead type. The overlays also indicate that the furthest "to the left" the critical points governing stability can be placed (for the curves representing $0.01 \leq \zeta \leq 1.0$) is along the circular arc segment 3'-7', since this arc is the image of the condition

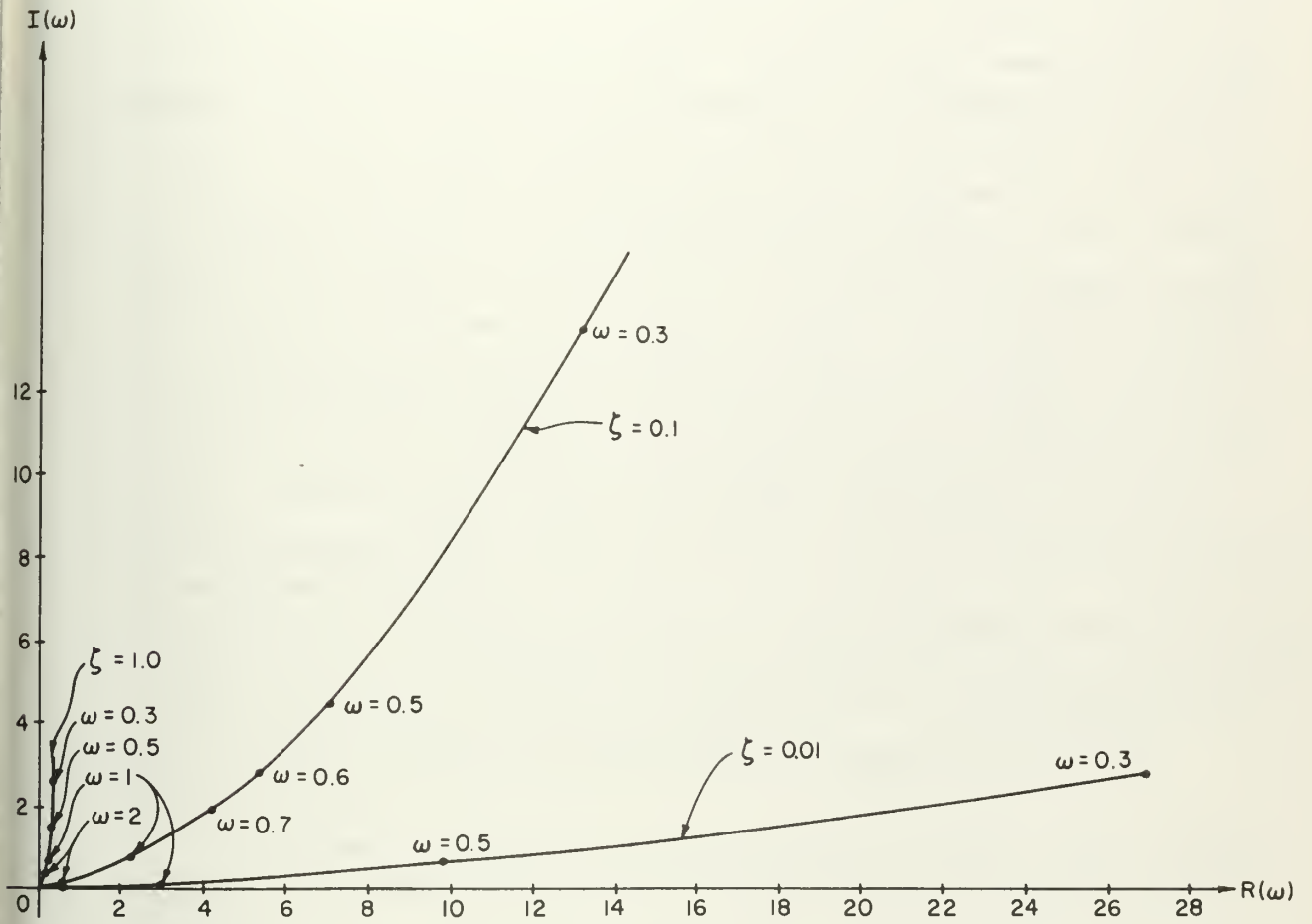


Fig. 6-10 : Controlled Element Plots, $-L_n$, for various ζ .

$0 \leq \alpha \leq 5.3$, $\beta = 0$. Also, in passing it may be noted that the condition $0 \leq \alpha \leq 5.3$, $\beta = 0.7$ is represented by arc 1'-5' (Fig. 4-18 indicates the definition of these points), i. e., β would play the role of T_N in the human-operator transfer function and the tiny hull formed by line 3'1', arc 1'-5', line 5'7' and arc 7'-3' would be of interest, (detail in Fig. 4-18). Since this region remains reasonably unchanged in the overlays for increasing frequency, it is possible to use the overlay for $\alpha = 1, 2$ or 3 , respectively, for an overall estimate of what is needed for α, β values. Furthermore, it is obvious that common values of α and β which provide stability over all ζ , can be found from those values of α and β yielding stability for the condition $\zeta = 0.01$, since it is this value which imposes the most severe stability requirement. Also, the plots suggest that increased gain might help the overall situation across the range of varying ζ , since the $\zeta = 1.0$ condition most assuredly benefits from increased gain, while the effects of increased gain upon the $\zeta = 0.01$ condition can hardly place that locus much closer to the real axis. Increased gain, however, may improve stability (for $\zeta = 0.01$) by shifting the frequency points further out along the locus and therefore further away from the region of criticality, Fig. 4-18 detail. Thus, values of $\alpha > 5.3$ only place the critical points depressingly close to the plant's Nyquist loci (all ζ), while for values of $\beta > 0.7$, a large portion of the circular segment for $0 \leq \alpha \leq 5.3$ lies in a region of instability for much of the range of $\zeta > 0.1$. The largest overall degree of stability across the varying range of ζ appears to be given by $\beta < 0.7$, $0 \leq \alpha \leq 5.3$.

It is reasonable, therefore, to begin the search for the α and β which meet specifications with starting values of $\alpha = 2$, $\beta = 0.5$. The specifications are taken to be

$$t_r = 2 \text{ seconds} = T_r$$

$$o_s = 15 \text{ percent} = O_s$$

$$t_s = 4 \text{ seconds} = T_s$$

Since the problem includes a search for the proper gain setting, K_c , this parameter is chosen to have a nominal value of unity, which represents the original conditions

for which this preliminary analysis has been conducted. It is noted that the plots suggest a final value of gain for K_c which is greater than unity, i. e., the Nyquist curves of Fig. 6-10 indicate that increased gain while not affecting the $\zeta = 0.01$ and $\zeta = 0.1$ curves substantially, ought to improve the response for the $\zeta = 1.0$ condition considerably. Thus, the starting values for the search are taken to be $\alpha = 2$, $\beta = 0.5$, $K_c = 1$.

Also,

$$f(K_c, \alpha, \beta) = (t_r - T_r)^2 + (o_s - O_s)^2 + (t_s - T_s)^2 \quad (6-13)$$

where

$$T_r = 2, O_s = 15, T_s = 4 \quad (6-14)$$

The step sizes chosen for the initial exploratory move are as follows:

$$\text{For } K_c: \text{ take step } \delta_1 = 0.2 \quad (6-15a)$$

$$\text{For } \alpha: \text{ take step } \delta_2 = 0.1 \quad (6-15b)$$

$$\text{For } \beta: \text{ take step } \delta_3 = 0.05 \quad (6-15c)$$

These exploratory increments have been chosen arbitrarily and, for explanatory purposes, they have been made distinct; obviously, they can also be set equal, if desired. The subscripts 1, 2, 3 are assigned to represent step increments along the K_c , α and β dimensional axes, respectively. Usually, the problem and the designer's experience suggest basic step sizes for each independent variable. During the termination procedure, the nominal step sizes of the increments are reduced until no further reduction is possible due either to solution indifference or equipment limitations.

Due to the pressure of time the search was conducted on analog-computer equipment rather than digital; the most satisfactory equipment for this type of problem is a hybrid computer containing both analog and digital capability, such equipment was not available. The computer was a Mid-Century Instrumatic (5068), a Mid-Century MC-300 six-channel recorder was used to obtain traces of system input signal, response, error, etc. The computation scheme is shown in

Fig. 6-11. If in Fig. 6-8, one lets

$$\frac{u}{e} = \frac{s\alpha + 1}{s\beta + 1} \quad (K_c = 1 \text{ temporarily}) \quad (6-16)$$

then

$$\beta u = \alpha e + (e - u) \frac{1}{s} \quad (6-17)$$

Also, if

$$\frac{u_1}{u} = \frac{1}{s} \quad (6-18)$$

then

$$\frac{v_1}{u_1} = \frac{\omega_n^2}{s + 2\zeta\omega_n} \quad (6-19)$$

and

$$v_1 = \frac{1}{s} \left[-2\zeta\omega_n v + \omega_n^2 u_1 \right] \quad (6-20)$$

Finally, if one sets

$$v = K_c v_1 \quad (6-21)$$

then Fig. 6-11 is a representation of the above relations.

It should be recognized that as a result of performing the search on analog equipment, the step-response characteristics (t_r, o_s, t_s) of the system were determined visually by reading the pen-recorder traces for $v(t)$. Once a near-optimal solution was obtained, it was not possible to reduce the increment sizes (δ) any further since any differences in the pen-recordings, for smaller δ values, were not distinguishable.

Generally, the difficulties in meeting the specifications are worse at the $\zeta = 1.0$ end of plant variability, i. e., worse for t_r and t_s , therefore, the search is started at that end. The initial base point for the search is designated, \bar{b}_1 , Fig. 6-12, the horizontal bar signifies the vector (multi-dimensional) nature of this point, i. e.,

$$\bar{b}_1 = (1, 2, 0.5) \quad (6-22)$$



Fig. 6-11: Analog Computer Setup for Direct Search Example

which is obtained from

$$K_c = 1, \alpha = 2, \beta = 0.5 \quad (6-23)$$

The step response of the system with a compensator which possesses such characteristics is

$$(t_r, o_s, t_s) = (4.8, 0, 9) \quad (6-24)$$

Therefore,

$$f(1, 2, 0.5) = [4.8-2]^2 + [0-15]^2 + [9-4]^2 = 257.81 \quad (6-25)$$

This value for the artificial function, f , is now compared with the value obtained at the next observation. The next observation is performed at the point $\bar{b}_1 + \bar{\delta}_1$, where $\bar{\delta}_1$ is a vector with all components zero except the one representing the independent variable, K_c . Thus,

$$\bar{b}_1 + \bar{\delta}_1 = (1.2, 2, 0.5) \quad (6-26)$$

since $|\bar{\delta}_1| = \delta_1 = 0.2$, from (6-15a). From these parameter values the response is found to be

$$(t_r, o_s, t_s) = (3.5, 0, 6.8) \quad (6-27)$$

Therefore,

$$f(1.2, 2, 0.5) = [3.5-2]^2 + [0-15]^2 + [6.8-4]^2 = 235.05 \quad (6-28)$$

In general, when there is little or no intuitive feel for a potential solution and one direction offers as much hope for success as another during the initial exploratory moves, the response for $\bar{b}_1 - \bar{\delta}_1$ is also obtained to determine if it yields better results than $\bar{b}_1 + \bar{\delta}_1$, or \bar{b}_1 . In this case, since

$$f(\bar{b}_1 + \bar{\delta}_1) < f(\bar{b}_1) \quad (6-29)$$

and, since it is recognized that a sluggish system ($\zeta = 1.0$) becomes even more

sluggish for reduced gain, the value of f produced by $\bar{b}_1 + \bar{\delta}_1$ is accepted as the smallest and $\bar{b}_1 + \bar{\delta}_1$ is designated the temporary base point, \bar{t}_{11} . The subscripts indicate this is the first temporary head (in a vector sense) and only one variable (K_c) has been perturbed. Thus, \bar{t}_{11} , represents the temporary base point about which the next exploratory move for the second independent variable, α , is made. To retain consistency of notation, let

$$\bar{t}_{10} = \bar{b}_1 \quad (6-30)$$

To find \bar{t}_{12} , i.e., the temporary head obtained from perturbing the second independent variable, the response for $\bar{t}_{11} + \bar{\delta}_2$ is first determined. Since $\bar{\delta}_2$ represents an increment to α of 0.1, the response for

$$\bar{t}_{11} + \bar{\delta}_2 = (1.2, 2.1, 0.5) \quad (6-31)$$

must be found. The performance index for these parameter values yields

$$(t_r, o_s, t_s) = (3.2, 0, 8) \quad (6-32)$$

Therefore,

$$f(1.2, 2.1, 0.5) = [3.2-2]^2 + [0-15]^2 + [8-4]^2 = 242.44 \quad (6-33)$$

The response for

$$\bar{t}_{11} - \bar{\delta}_2 = (1.2, 1.9, 0.5) \quad (6-34)$$

is now found. The performance index becomes

$$(t_r, o_s, t_s) = (3, 0, 6.8) \quad (6-35)$$

Therefore,

$$f(1.2, 1.9, 0.5) = [3-2]^2 + [0-15]^2 + [6.8-4]^2 = 233.8 \quad (6-36)$$

Since

$$f(\bar{t}_{11} - \bar{\delta}_2) < f(\bar{t}_{11}) < f(\bar{t}_{11} + \bar{\delta}_2) \quad (6-37)$$

the next temporary head is taken to be

$$\bar{t}_{12} = \bar{t}_{11} - \bar{\delta}_2$$

To find, \bar{t}_{13} , the third and last (in this example) independent variable, β , is perturbed by $\bar{\delta}_3$. First, the response for

$$(\bar{t}_{12} + \bar{\delta}_3) = (1.2, 1.9, 0.55) \quad (6-38)$$

is found.

The performance index yields

$$(t_r, o_s, t_s) = (3.8, 0, 5.8) \quad (6-39)$$

Therefore,

$$f(1.2, 1.9, 0.55) = [3.8-2]^2 + [0-15]^2 + [5.8-4]^2 = 231.48 \quad (6-40)$$

The response for

$$\bar{t}_{12} - \bar{\delta}_3 = (1.2, 1.9, 0.45) \quad (6-41)$$

is now found. The performance index yields

$$(t_r, o_s, t_s) = (3.9, 0, 5.5) \quad (6-42)$$

Therefore,

$$f(1.2, 1.9, 0.45) = [3.9-2]^2 + [0-15]^2 + [5.5-4]^2 = 230.86 \quad (6-43)$$

Since,

$$f(\bar{t}_{12} - \bar{\delta}_3) < f(\bar{t}_{12} + \bar{\delta}_3) < f(\bar{t}_{12}) \quad (6-44)$$

the next temporary head is taken to be

$$\bar{t}_{13} = \bar{t}_{12} - \bar{\delta}_3 \longrightarrow \bar{b}_2 \quad (6-45)$$

All the independent variables have now been perturbed; the last temporary head, \bar{t}_{13} , is taken to be the new, or second, base point, \bar{b}_2 .

With the establishment of the second base point, \bar{b}_2 , the pattern strategy is brought to bear. The pattern move uses the information generated in the exploratory moves which, initially, establish the pattern and later modify the pattern, as

the example is to demonstrate. The pattern moves are formulated upon the philosophy of "nothing succeeds like success". Based upon the success in the past, of motion to \bar{b}_2 from \bar{b}_1 , the assumption is made that continued motion in the same direction remains the most desirable course of action. Thus, the new temporary head for the second set of exploratory moves is obtained by doubling the magnitude of $(\bar{b}_2 - \bar{b}_1)$ and proceeding in the same direction. Hence,

$$\bar{t}_{20} = \bar{b}_1 + 2(\bar{b}_2 - \bar{b}_1) = \bar{b}_2 + (\bar{b}_2 - \bar{b}_1) = 2\bar{b}_2 - \bar{b}_1 \quad (6-46)$$

Therefore, from (6-22) and (6-45)

$$\bar{t}_{20} = 2\bar{b}_2 - \bar{b}_1 = 2(1.2, 1.9, 0.45) - (1, 2, 0.5) \quad (6-47)$$

$$= (2.4, 3.8, 0.9) - (1, 2, 0.5) \\ = (1.4, 1.8, 0.4) \quad (6-48)$$

The performance index becomes

$$(t_r, o_s, t_s) = (4, 0, 4.2) \quad (6-49)$$

Therefore,

$$f(1.4, 1.8, 0.4) = [4-2]^2 + [0-15]^2 + [4.2-4]^2 = 229.04 \quad (6-50)$$

The response for

$$\bar{t}_{20} + \bar{\delta}_1 = (1.4 + 0.2, 1.8, 0.4) = (1.6, 1.8, 0.4) \quad (6-51)$$

is now found. The performance index yields

$$(t_r, o_s, t_s) = (3, 0, 4.8) \quad (6-52)$$

Therefore,

$$f(1.6, 1.8, 0.4) = [3-2]^2 + [0-15]^2 + [4.8-4]^2 = 226.4 \quad (6-53)$$

It is found that

$$f(\bar{t}_{20} + \bar{\delta}_1) < f(\bar{t}_{20}) \quad (6-54)$$

If one reasons as previously with regard to the gain, K_c , the calculation for $\bar{t}_{20} - \bar{\delta}_1$ can be skipped and \bar{t}_{21} can be taken to be

$$\bar{t}_{21} = \bar{t}_{20} + \bar{\delta}_1 \quad (6-55)$$

The search for \bar{t}_{22} now begins. To this end, the response for $\bar{t}_{21} + \bar{\delta}_2$ is found, where

$$\bar{t}_{21} + \bar{\delta}_2 = (1.6, 1.9, 0.4) \quad (6-56)$$

The performance index yields

$$(t_r, o_s, t_s) = (3.1, 0, 4.5) \quad (6-57)$$

Therefore,

$$f(1.6, 1.9, 0.4) = [3.1-2]^2 + [0-15]^2 + [4.5-4]^2 = 226.46 \quad (6-58)$$

The response for

$$\bar{t}_{21} - \bar{\delta}_2 = (1.6, 1.7, 0.4) \quad (6-59)$$

yields

$$(t_r, o_s, t_s) = (3.0, 0, 4.0) \quad (6-60)$$

Therefore,

$$f(1.6, 1.7, 0.4) = [3-2]^2 + [0-15]^2 + [4-4]^2 = 226 \quad (6-61)$$

Again, since

$$f(\bar{t}_{21} - \bar{\delta}_2) < f(\bar{t}_{21} + \bar{\delta}_2) < f(\bar{t}_{21}) \quad (6-62)$$

the next temporary head is taken to be

$$\bar{t}_{22} = \bar{t}_{21} - \bar{\delta}_2 \quad (6-63)$$

To find, \bar{t}_{23} , the last independent variable, β , is perturbed by δ_3 . First, the response is found for

$$\bar{t}_{22} + \bar{\delta}_3 = (1.6, 1.7, 0.45) \quad (6-64)$$

The performance index yields

$$(t_r, o_s, t_s) = (3.8, 0, 6.0) \quad (6-65)$$

Therefore,

$$f(1.6, 1.7, 0.45) = [3.8-2]^2 + [0-15]^2 + [6-4]^2 = 232.24 \quad (6-66)$$

The response is found for

$$\bar{t}_{22} - \bar{\delta}_3 = (1.6, 1.7, 0.35) \quad (6-67)$$

The performance index yields

$$(t_r, o_s, t_s) = (2.8, 0, 4.3) \quad (6-68)$$

Therefore,

$$f(1.6, 1.7, 0.35) = [2.8-2]^2 + [0-15]^2 + [4.3-4]^2 = 225.73 \quad (6-69)$$

Since

$$f(\bar{t}_{22} - \bar{\delta}_3) < f(\bar{t}_{22}) < f(\bar{t}_{22} + \bar{\delta}_3) \quad (6-70)$$

the next temporary head is taken to be

$$\bar{t}_{23} = \bar{t}_{22} - \bar{\delta}_3 \longrightarrow \bar{b}_3 \quad (6-71)$$

Since all the independent variables have again been perturbed the last temporary head, \bar{t}_{23} , is taken as the new, or third, base point, \bar{b}_3 . The process is continued; from (6-45) and (6-71) is found

$$\bar{t}_{30} = 2\bar{b}_3 - \bar{b}_2 \quad (6-72)$$

$$\begin{aligned} &= 2(1.6, 1.7, 0.35) - (1.2, 1.9, 0.45) \\ &= (3.2, 3.4, 0.7) - (1.2, 1.9, 0.45) \\ &= (2.0, 1.5, 0.25) \end{aligned} \quad (6-73)$$

It is noted that repeated success in the same direction, for all the independent variables, has caused the pattern to grow. Thus, although only two iterations have been conducted, the gain value, K_c , is at 2.0 after starting at 1.0 and being constrained to basic step sizes $\delta_1 = 0.2$.

The performance index becomes for \bar{t}_{30} ,

$$(t_r, o_s, t_s) = (2.0, 0, 3.6) \quad (6-74)$$

Therefore,

$$f(2, 1.5, 0.25) = [2-2]^2 + [0-15]^2 + [3.6-4]^2 = 225.16 \quad (6-75)$$

It is noted that the set of values

$$(K_c, \alpha, \beta) = (2.0, 1.5, 0.25) \quad (6-76)$$

is the first triple for which the specifications are met by all three elements of the vector performance index (t_r, o_s, t_s) . The results of the search to this point are encouraging. The artificial function f which has served the purpose of finding a surface of acceptable performance no longer provides a suitable comparison test; the magnitudes of the individual differences

$$|t_r - T_r|, |o_s - O_s|, |t_s - T_s| \quad (6-77)$$

are monitored in an attempt to distinguish "best" performance. The response for

$$\bar{t}_{30} + \bar{\delta}_1 = (2.2, 1.5, 0.25) \quad (6-78)$$

yields a performance index

$$(t_r, o_s, t_s) = (1.7, 0, 3.8) \quad (6-79)$$

Therefore,

$$f(2.2, 1.5, 0.25) = [1.7-2]^2 + [0-15]^2 + [3.8-4]^2 = 225.13 \quad (6-80)$$

The response is still without overshoot, so the exploratory move $(\bar{t}_{30} - \bar{\delta}_1)$ is again omitted and

$$\bar{t}_{31} = \bar{t}_{30} + \bar{\delta}_1 \quad (6-81)$$

The choice has been made, not because

$$f(\bar{t}_{30} + \bar{\delta}_1) < f(\bar{t}_{30}) \quad (6-82)$$

although that is true, but because the response for $\bar{t}_{30} + \bar{\delta}_1$ is further inside specifications than that for \bar{t}_{30} . This is discussed in detail below.

The search for \bar{t}_{32} is begun. First, the response for

$$\bar{t}_{31} + \bar{\delta}_2 = (2.2, 1.6, 0.25) \quad (6-83)$$

is found. The performance index gives

$$(t_r, o_s, t_s) = (2, 0, 4) \quad (6-84)$$

Therefore,

$$f(2.2, 1.6, 0.25) = [2-2]^2 + [0-15]^2 + [4-4]^2 = 225 \quad (6-85)$$

For

$$\bar{t}_{31} - \bar{\delta}_2 = (2.2, 1.4, 0.25) \quad (6-86)$$

the performance index yields

$$(t_r, o_s, t_s) = (2, 0, 3.5) \quad (6-87)$$

At this point, the investigator faces a quandary. With the boundary of acceptable performance pierced, and the artificial function f effectively discarded, how is the decision regarding the "best" temporary base point to be made? The three possibilities, \bar{t}_{30} , $\bar{t}_{30} + \bar{\delta}_1$ and $\bar{t}_{30} - \bar{\delta}_1$ yield performance indices (1.7, 0, 3.8), (2, 0, 4) and (2, 0, 3.5), respectively. All are within specifications; which is best? While $\bar{t}_{30} + \bar{\delta}_1$ yields a set just meeting requirements, and on that basis could be discarded, how can a decision be made between \bar{t}_{30} and $\bar{t}_{30} - \bar{\delta}_1$? Obviously, what the one gains on rise time the other gains on settling time. Two possible techniques suggest themselves: a) a coin is tossed, i.e., one invokes the Principle of Insufficient Reason, b) the specifications are reduced; a new artificial function, f_1 , is formed, and another comparison measure of which set (K_c, α, β) is "best" is thereby provided.

To this end, a new set of requirements

$$T_r = 1.5, \quad O_s = 15, \quad T_s = 3.0 \quad (6-88)$$

are specified; it should be noted that rise time and settling time have been retained at the same ratio as previously. Under these circumstances,

$$f_1(\bar{t}_{31}) = f_1(2.2, 1.5, 0.25) = [1.7-1.5]^2 + [0-15]^2 + [3.8-3]^2 = 225.68 \quad (6-89)$$

$$f_1(\bar{t}_{31} + \bar{\delta}_2) = f_1(2.2, 1.6, 0.25) = [2-1.5]^2 + [0-15]^2 + [4-3]^2 = 226.25 \quad (6-90)$$

$$f_1(\bar{t}_{31} - \bar{\delta}_2) = f_1(2.2, 1.4, 0.25) = [2-1.5]^2 + [0-15]^2 + [3.5-3]^2 = 225.50 \quad (6-91)$$

Therefore, since

$$f_1(\bar{t}_{31} - \bar{\delta}_2) < f_1(\bar{t}_{31}) < f_1(\bar{t}_{31} + \bar{\delta}_2) \quad (6-92)$$

the next temporary head is taken to be

$$\bar{t}_{32} = \bar{t}_{31} - \bar{\delta}_2 \quad (6-93)$$

To find, \bar{t}_{33} , the variable, δ , is augmented by $\bar{\delta}_3$. First, the response for

$$\bar{t}_{32} + \bar{\delta}_3 = (2.2, 1.4, 0.30) \quad (6-94)$$

is found. The performance index yields

$$(t_r, o_s, t_s) = (2, 0, 3.5) \quad (6-95)$$

Therefore,

$$f_1(2.2, 1.4, 0.30) = [2-1.5]^2 + [0-15]^2 + [3.5-3]^2 = 225.50 \quad (6-96)$$

The response for

$$\bar{t}_{32} - \bar{\delta}_3 = (2.2, 1.4, 0.2) \quad (6-97)$$

is found. The performance index yields

$$(t_r, o_s, t_s) = (1.9, 0, 3.4) \quad (6-98)$$

Therefore,

$$f_1(2.2, 1.4, 0.2) = [1.9-1.5]^2 + [0-15]^2 + [3.4-3]^2 = 225.32 \quad (6-99)$$

Since

$$f_1(\bar{t}_{32} - \bar{\delta}_3) < f_1(\bar{t}_{32}) = f_1(\bar{t}_{32} + \bar{\delta}_3) \quad (6-100)$$

the next temporary head is taken to be

$$\bar{t}_{33} = \bar{t}_{32} - \bar{\delta}_3 \longrightarrow \bar{b}_4 \quad (6-101)$$

Since all the independent variables have again been perturbed, the last temporary head, \bar{t}_{33} , is taken as the new, or fourth, base point, \bar{b}_4 . From (6-71) and (6-101) is calculated

$$\bar{t}_{40} = 2\bar{b}_4 - \bar{b}_3 \quad (6-102)$$

$$\begin{aligned} &= 2(2.2, 1.4, 0.2) - (1.6, 1.7, 0.35) \\ &= (4.4, 2.8, 0.4) - (1.6, 1.7, 0.35) \\ &= (2.8, 1.1, 0.05) \end{aligned} \quad (6-103)$$

The response for, \bar{t}_{40} , is determined and the characteristics yield

$$(t_r, o_s, t_s) = (1.4, 0, 2.75) \quad (6-104)$$

Therefore,

$$f_1(2.8, 1.1, 0.05) = [1.4-1.5]^2 + [0-15]^2 + [2.75-3]^2 \quad (6-105)$$

The new function, f_1 , must now be discarded since the set of values for K_c, α, β , (6-103), provides a response which again is within the specifications on t_r, o_s, t_s . Thus, a second new comparison function is formed for the changed requirements

$$T_r = 1.0, O_s = 10, T_s = 2 \quad (6-106)$$

Therefore,

$$f_2(2.8, 1.1, 0.05) = [1.4-1]^2 + [0-10]^2 + [2.75-2]^2 = 100.72, \quad (6-107)$$

The response for

$$\bar{t}_{40} + \bar{\delta}_1 = (3.0, 1.1, 0.05) \quad (6-108)$$

is now found. The performance index yields

$$(t_r, o_s, t_s) = (1.4, 0, 2.3) \quad (6-109)$$

Therefore,

$$f_2(3.0, 1.1, 0.05) = [1.4-1]^2 + [0-10]^2 + [2.3-2]^2 = 100.25 \quad (6-110)$$

Since, the response is again without overshoot, the exploratory move $(\bar{t}_{30} - \bar{\delta}_1)$ is again omitted. Also,

$$f_2(\bar{t}_{40} + \bar{\delta}_1) < f_2(\bar{t}_{40}) \quad (6-111)$$

so

$$\bar{t}_{41} = \bar{t}_{40} + \bar{\delta}_1 \quad (6-112)$$

The search for, \bar{t}_{42} , is begun. First, the response is found for

$$\bar{t}_{41} + \bar{\delta}_2 = (3, 1.2, 0.05) \quad (6-113)$$

The performance index is

$$(t_r, o_s, t_s) = (2, 0, 2.9) \quad (6-114)$$

Therefore,

$$f_2(3, 1.2, 0.05) = [2-1]^2 + [0-10]^2 + [2.9-2]^2 = 101.81 \quad (6-115)$$

For

$$\bar{t}_{41} - \bar{\delta}_2 = (3, 0, 1.0, 0.05) \quad (6-116)$$

the performance index gives

$$(t_r, o_s, t_s) = (1.7, 0, 2.4) \quad (6-117)$$

Therefore,

$$f_2(\bar{t}_{41} - \bar{\delta}_2) = f_2(3, 1, 0.05) = [1.7-1]^2 + [0-10]^2 + [2.4-2]^2 = 100.65 \quad (6-118)$$

Thus, since

$$f_2(\bar{t}_{41}) < f_2(\bar{t}_{41} - \bar{\delta}_2) < f_2(\bar{t}_{41} + \bar{\delta}_2) \quad (6-119)$$

the next temporary head is taken to be

$$\bar{t}_{42} = \bar{t}_{41} \quad (6-120)$$

To find, \bar{t}_{43} , the variable, β , is increased by $\bar{\delta}_3$. First, the response is found for

$$\bar{t}_{42} + \bar{\delta}_3 = (3, 0, 1.1, 0.1) \quad (6-121)$$

The performance index yields

$$(t_r, o_s, t_s) = (1.8, 0, 3.2) \quad (6-122)$$

Therefore,

$$f_2(3, 1.1, 0.1) = [1.8-1]^2 + [0-10]^2 + [3.2-2]^2 = 102.08 \quad (6-123)$$

The response for

$$\bar{t}_{42} - \bar{\delta}_3 = (3, 0, 1.1, 0) \quad (6-124)$$

cannot be used due to practical limitations on the analog computer, i.e., due primarily to the layout used to solve the problem. An even more important justification for not reducing β to zero, however, is that it is not possible to realize such a compensation network, within the given class to which the compensator design is constrained. Furthermore, since the artificial function has been reduced twice already, the design is well within original specifications, for $\zeta = 1.0$. Thus, since

$$f_2(\bar{t}_{42}) < f_2(\bar{t}_{42} + \bar{\delta}_3) \quad (6-125)$$

$$\text{i.e., } f_2(\bar{t}_{42}) = f_2(\bar{t}_{41}) = f_2(\bar{t}_{40} + \bar{\delta}_1) < f_2(\bar{t}_{42} + \bar{\delta}_3) \quad (6-126)$$

$$100.25 < 102.08$$

the next temporary head is taken to be

$$\bar{t}_{43} = \bar{t}_{42} \longrightarrow \bar{b}_5 \quad (6-127)$$

Since all the independent variables have again been perturbed, the last temporary head, \bar{t}_{43} , is taken as the new, or fifth, basepoint, \bar{b}_5 . From (6-97) and (6-108) is calculated

$$\bar{t}_{50} = 2\bar{b}_5 - \bar{b}_4 \quad (6-128)$$

$$\begin{aligned} &= 2(3, 1.1, 0.05) - (2.2, 1.4, 0.2) \\ &= (6, 2.2, 0.1) - (2.2, 1.4, 0.2) \\ &= (3.8, 0.8, -0.1) \longrightarrow (3.8, 0.8, 0.05) \end{aligned} \quad (6-129)$$

That is, under the practical constraint that β not be less than 0.05, β is set equal to 0.05 whenever the process indicates $\beta < 0.05$.

The response for

$$\bar{t}_{50} = (3.8, 0.8, 0.05) \quad (6-130)$$

is determined. The performance index yields

$$(t_r, o_s, t_s) = (1.4, 0, 2.5) \quad (6-131)$$

Therefore,

$$f_2(3.8, 0.8, 0.05) = [1.4-1]^2 + [0-10]^2 + [2.5-2]^2 = 100.41 \quad (6-132)$$

The response is now found for

$$\bar{t}_{50} + \bar{\delta}_1 = (4.0, 0.8, 0.05) \quad (6-133)$$

The performance index yields

$$(t_r, o_s, t_s) = (1.2, 0, 2.0) \quad (6-134)$$

Therefore,

$$f_2(4, 0.8, 0.05) = [1.2-1]^2 + [0-10]^2 + [2-2]^2 = 100.04 \quad (6-135)$$

Again, the exploratory move, $\bar{t}_{50} - \bar{\delta}_1$, is omitted, and since

$$f_2(\bar{t}_{50} + \bar{\delta}_1) < f_2(\bar{t}_{50}) \quad (6-136)$$

$$\bar{t}_{51} = \bar{t}_{50} + \bar{\delta}_1$$

The search for, \bar{t}_{52} , is begun. First, the response is found for

$$\bar{t}_{51} + \bar{\delta}_2 = (4, 0.9, 0.05) \quad (6-137)$$

The performance index gives

$$(t_r, o_s, t_s) = (1.1, 0, 2.0) \quad (6-138)$$

Therefore,

$$f_2(4, 0.9, 0.05) = [1.1-1]^2 + [0-10]^2 + [2-2]^2 = 100.01 \quad (6-139)$$

For,

$$\bar{t}_{51} - \bar{\delta}_2 = (4, 0.7, 0.05) \quad (6-140)$$

the performance index yields

$$(t_r, o_s, t_s) = (1.3, 0, 2.5) \quad (6-141)$$

Therefore,

$$f_2(4, 0.7, 0.05) = [1.3-1]^2 + [0-10]^2 + [2.5-2]^2 = 100.34 \quad (6-142)$$

Thus, since

$$f_2(\bar{t}_{51} + \bar{\delta}_2) < f_2(\bar{t}_{51}) < f_2(\bar{t}_{51} - \bar{\delta}_2) \quad (6-143)$$

the next temporary head is taken to be

$$\bar{t}_{52} = \bar{t}_{51} + \bar{\delta}_2 \quad (6-144)$$

To find, \bar{t}_{53} , the variable, β , is increased by δ_3 . First, the response is found for

$$\bar{t}_{52} + \bar{\delta}_3 = (4, 0.9, 0.1) \quad (6-145)$$

The performance index yields

$$(t_r, o_s, t_s) = (1.2, 0, 2.2) \quad (6-146)$$

Therefore,

$$f_2(4, 0.9, 0.1) = [1.2 - 1]^2 + [0 - 10]^2 + [2.2 - 2]^2 = 100.08 \quad (6-147)$$

Again, the response for $\bar{t}_{52} - \bar{\delta}_3$ cannot be found so the $\bar{t}_{52} + \bar{\delta}_3$ response is matched against the \bar{t}_{52} response and it is found that

$$f(\bar{t}_{52}) < f(\bar{t}_{52} + \bar{\delta}_3) \quad (6-148)$$

Therefore, the next temporary head is taken to be

$$\bar{t}_{53} = \bar{t}_{52} \longrightarrow \bar{b}_6 \quad (6-149)$$

Since all the independent variables have again been perturbed, the last temporary head, \bar{t}_{53} , is taken as the new, or sixth, base point, \bar{b}_6 . From (6-127) and (6-149) is found

$$\bar{t}_{60} = 2\bar{b}_6 - \bar{b}_5 \quad (6-150)$$

$$= 2(4, 0.9, 0.05) - (3, 1.1, 0.05)$$

$$= (8, 1.8, 0.1) - (3, 1.1, 0.05)$$

$$= (5, 0.7, 0.05) \quad (6-151)$$

The response is determined for

$$\bar{t}_{60} = (5, 0.7, 0.05) \quad (6-152)$$

The performance index yields

$$(t_r, o_s, t_s) = (1.1, 0, 2.1) \quad (6-153)$$

Therefore,

$$f_2(5, 0.7, 0.05) = [1.1-1]^2 + [0-10]^2 + [2.1-2]^2 = 100.02 \quad (6-154)$$

The response for

$$\bar{t}_{60} + \bar{\delta}_1 = (5.2, 0.7, 0.05) \quad (6-155)$$

is now found.

The performance index yields

$$(t_r, o_s, t_s) = (1.1, 0, 2.2) \quad (6-156)$$

Therefore,

$$f_2(5.2, 0.7, 0.05) = [1.1-1]^2 + [0-10]^2 + [2.2-2]^2 = 100.05 \quad (6-157)$$

The response for

$$\bar{t}_{60} - \bar{\delta}_1 = (4.8, 0.7, 0.05) \quad (6-158)$$

if found. The performance index yields

$$(t_r, o_s, t_s) = (1.1, 0, 2.2) \quad (6-159)$$

Therefore,

$$f_2(4.8, 0.7, 0.05) = [1.1-1]^2 + [0-10]^2 + [2.2-2]^2 = 100.05 \quad (6-160)$$

Since

$$f_2(\bar{t}_{60}) \sim f_2(\bar{t}_{60} + \bar{\delta}_1) = f_2(\bar{t}_{60} - \bar{\delta}_1) \quad (6-161)$$

the next temporary head (basepoint) is chosen to be

$$\bar{t}_{61} = \bar{t}_{60} \quad (6-162)$$

The search is started for \bar{t}_{62} ; the response is found for

$$\bar{t}_{61} + \bar{\delta}_2 = (5, 0.8, 0.05) \quad (6-163)$$

The performance index measures

$$(t_r, o_s, t_s) = (1.05, 0, 2) \quad (6-164)$$

Therefore,

$$f_2(5, 0.8, 0.05) = [1.05-1]^2 + [0-10]^2 + [2-2]^2 = 100.0025 \quad (6-165)$$

For,

$$\bar{t}_{61} - \bar{\delta}_2 = (5, 0.6, 0.05) \quad (6-166)$$

the performance index yields

$$(t_r, o_s, t_s) = (1.1, 0, 2.0) \quad (6-167)$$

Therefore,

$$f_2(5, 0.6, 0.05) = [1.1-1]^2 + [0-10]^2 + [2-2]^2 = 100.01 \quad (6-168)$$

Since,

$$f_2(\bar{t}_{61} + \bar{\delta}_2) < f_2(\bar{t}_{61} - \bar{\delta}_2) < f_2(\bar{t}_{61}) \quad (6-169)$$

the temporary head is taken to be

$$\bar{t}_{62} = \bar{t}_{61} + \bar{\delta}_2 \quad (6-170)$$

To find, \bar{t}_{63} , the third independent variable, β , is increased by, δ_3 , and the response is found for

$$\bar{t}_{62} + \bar{\delta}_3 = (5, 0.8, 0.1) \quad (6-171)$$

The performance index yields

$$(t_r, o_s, t_s) = (1, 0, 2) \quad (6-172)$$

Therefore,

$$f_2(5, 0.8, 0.1) = [1-1]^2 + [0-10]^2 + [2-2]^2 = 100 \quad (6-173)$$

Again, the response for $\bar{t}_{62} - \bar{\delta}_3$ cannot be found so the $\bar{t}_{62} + \bar{\delta}_3$ response is matched against the \bar{t}_{62} response and it is found that

$$f(\bar{t}_{62} + \bar{\delta}_3) < f(\bar{t}_{62}) \quad (6-174)$$

Therefore, the next temporary head is taken to be

$$\bar{t}_{63} = \bar{t}_{62} + \bar{\delta}_3 \longrightarrow \bar{b}_7 \quad (6-175)$$

Since all the independent variables have been perturbed, the last temporary head, \bar{t}_{63} , is taken as the new, or seventh, basepoint, \bar{b}_7 . Thus, from (6-149) and (6-175)

$$\bar{t}_{70} = 2\bar{b}_7 - \bar{b}_6 \quad (6-176)$$

$$\begin{aligned} &= 2(5, 0.8, 0.1) - (4, 0.9, 0.05) \\ &= (10, 1.6, 0.2) - (4, 0.9, 0.05) \\ &= (6, 0.7, 0.15) \end{aligned} \quad (6-177)$$

The response for

$$\bar{t}_{70} = (6, 0.7, 0.15) \quad (6-178)$$

can be determined, etc. The process can be continued, thus providing much additional information regarding the correlation between (K_c, α, β) space and the space of (t_r, o_s, t_s) . A solution providing satisfactory response across the range of varying ζ is $(K_c, \alpha, \beta) = (5, 1.0, 0.05)$. The values of the performance index for various ζ are displayed in Table 6-1. It is clear from Table 6-1 that, $K_c = 5$, $\alpha = 1.0$, and $\beta = 0.05$ form a satisfactory compensator which meets the original specifications.

TABLE 6-1

| K_c | $\alpha = 1.0$ | $\beta = 0.05$ | ζ | t_r | o_s | t_s |
|-------|----------------|----------------|---------|-------|-------|-------|
| | | | 1.0 | 1 | 0 | 2 |
| | | | 0.9 | 0.95 | 0 | 1.95 |
| | | | 0.8 | 0.8 | 0 | 1.6 |
| | | | 0.7 | 0.5 | 0 | 1.3 |
| | | | 0.6 | 0.3 | 0 | 1.0 |
| | | | 0.5 | 0.25 | 0 | 0.9 |
| | | | 0.4 | 0.2 | 0 | 0.25 |
| | | | 0.3 | 0.17 | 0 | 0.20 |
| | | | 0.2 | 0.15 | 2 | 0.17 |
| | | | 0.1 | 0.1 | 4 | 0.15 |
| | | | 0.05 | 0.1 | 7 | 0.5 |
| | | | 0.01 | 0.1 | 9 | 0.9 |

$$T_r = 2, O_s = 15, T_s = 4 \quad (6-14)$$

for all ζ in the range, $0.01 \leq \zeta \leq 1.0$. Indeed, this compensator meets the more stringent requirements of

$$T_r = 1, O_s = 10, T_s = 2 \quad (6-106)$$

across the entire range specified for varying ζ .

A review of the example indicates how quickly the search converges upon a solution; a digital computation program can, of course, provide the answer obtained with considerably more speed and accuracy. The addition of the delay time (pure lag), which is the major change to the example, introduced by the human operator, in no way modifies the search procedure used. The delay interval has its effect upon the system response characteristics and may reduce the prospect of success in finding a "best-overall" compensator of a given class but the direct search method with pattern strategy retains its general applicability.

It should be noted again that, in the example just offered, the human-operator transfer function was set equal to unity and the variable-parameter-mapping plots used to estimate the approximate values that α and β must have if compensation is to achieve the purpose prescribed. Thus, the three-variable mappings were allowed to degenerate to a two-variable representation and the values of α and β in this instance, had a direct effect upon the location of the critical points governing stability. Normally, when the Ω mappings are used to represent the human-operator adaptive range the influence of α and β values upon system stability is not so easily discernible. In such cases, the selection of values for α and β influences the shape of the $-L_n$ curve, which represents the system dynamics; the manner in which this curve passes through the Ω regions is, of course, what ultimately determines overall (with a man in the loop) system stability and performance. Thus, in an analysis with a human operator in the system, the effects of the K_c, α and β values are only indirectly observable before they are selected. Upon selection, however, if the compensated $-L_n$ is plotted the effects are clearly visible. This situation is no different than that routinely faced by the control engineer when dealing with less animate systems. Experience is quickly gained concerning the results of introducing particular com-

pensator configurations (picking values for α and β); to this end, with a definite goal for the shape of the modified $-L_n$ curve the Bode' diagram (more quickly plotted than a Nyquist curve) can be used to advantage.

6.7 Summary

A discussion of scalar performance indices and some of their limitations is followed by a laconic survey of methods for solving the "best" performance problem, especially for varying-parameter plants. The notion of a "vector" performance index is introduced and a technique for establishing one response as "better" than another is presented. This technique rests intimately upon the formulation of an "artificial" function which when reduced to zero identically satisfies the upper bounds on system requirements. These requirements are set forth in the "vector" index. The design for a "best-overall" compensator to meet specifications upon a varying-parameter plant is obtained by using a "direct search" method (especially suited to the electronic computer). This search method incorporates a "pattern strategy" to aid in finding the next set of trial values for the compensator; the strategy affords rapid convergence to a solution as is demonstrated by an example.

Finally, the example itself, displays the versatility of the variable-parameter mappings presented in Chapters Two-Five. By setting the human-operator transfer function equal to unity, it is shown that the mappings have use for "inanimate" system design.

Empirical justification of the conclusions from this theory (with a human operator in the loop) must wait upon additional research and the publication of experimental results issuing from deterministic inputs driving controlled elements other than pure gains. It is recognized that the bulk of the evidence justifying the form of human-operator transfer function used in this study has been accumulated for systems driven by stochastic input signals. Still, the initial range selection for the parameters K_h , τ , T_L , T_I and T_N included the effects of calculations published for deterministic inputs and modifications, if needed, are easily incorporated. With a technique provided for comparing one multi-dimensional index against another, the methods of this chapter offer the prospect for more realistic performance measures than the standard "figure of merit" of scalar variety affords.

APPENDIX A

A.1 Human Operator Zero Locus

It is desired to establish (4-13) as the equation of a circle. The transformation (A-1) is to be considered

$$w = u + jv = \frac{a + jb}{1 + x + jy} \quad (\text{A-1})$$

Of interest are the conditions $x = 0$, $y = xT_L$, thus

$$w = \frac{a + jb}{(1 + x) + jy} \cdot \frac{(1 + x) - jy}{(1 + x) - jy}$$

$$w = \frac{a(1 + x) + by + j[b(1 + x) - ay]}{(1 + x)^2 + y^2} \quad (\text{A-2})$$

The real and imaginary parts of each side of (A-2) are set equal

$$u = \frac{a(1 + x) + by}{(1 + x)^2 + y^2} \quad (\text{A-3})$$

$$v = \frac{b(1 + x) - ay}{(1 + x)^2 + y^2} \quad (\text{A-4})$$

Similarly,

$$1 + x + jy = \frac{a + jb}{u + jv} \cdot \frac{u - jv}{u - jv}$$

$$1 + x + jy = \frac{au + bv + j(bu - av)}{u^2 + v^2}$$

therefore

$$1+x = \frac{au+bv}{u^2+v^2} \quad ; \quad (\text{A-5})$$

$$y = \frac{bu-av}{u^2+v^2} \quad (\text{A-6})$$

For $x=0$, (A-5) gives the locus in the space of the mapping and (A-6) permits recovery of T_L for any given t

$$\therefore u^2+v^2-au-bv=0 \quad (\text{A-7})$$

If the square in (A-7) is completed, one obtains the familiar equation of a circle

$$\left(u - \frac{a}{2}\right)^2 + \left(v - \frac{b}{2}\right)^2 = \left[\frac{\sqrt{a^2+b^2}}{2}\right]^2 \quad (\text{A-8})$$

APPENDIX B

B.1 The Parabolic Effect

A typical plot of a face from (T_L, T_I, T_N) space, T_L constant, appears in Fig. B-1. Without loss of generality, a set of xy axes is shown so oriented as to simplify the following exposition. Of interest is the complete mapping corresponding to face ODCEO in Fig. B-2. This mapping can be found by allowing a typical element of this face such as $\alpha\beta$ in Fig. B-2 to scan from line OD to line EC. The effect, as has been seen in Chapter Four, is a plot similar to that shown in Fig. B-1, except for translation of axes. When α is halfway down line DC, β is halfway down OE, Fig. B-2, and similarly in the mapping space, Fig. B-1, i.e., linearity holds through the mapping as is proved later in this Appendix. Thus, α' is also halfway down the line $D'C'$ of Fig. B-1 as is β' halfway along $O'E'$. Each of these points is given arbitrary coordinate designation in Fig. B-1 and it is noted

$$x_3 = x_4 = x_5 = 0$$

for this choice of axes. The mapping of any element $\alpha\beta$ is obtained by joining the points $\alpha'\beta'$ in Fig. B-1. As shown in Fig. B-3 the family of these lines extends outside the end point mapping of Fig. B-1. It is shown that the envelope of this family of lines $\alpha'\beta'$ is a parabolic arc as suggested by the many dotted lines of this family drawn in Fig. B-3. For the geometry of Fig. B-1

$$\beta'_x = 0$$

$$\beta'_y = \lambda(y_3 - y_5) = \lambda y_3$$

$$\alpha'_x = \lambda(x_2 - x_4) = \lambda x_2$$

$$\alpha'_y = \lambda(y_2 - y_4) + y_4$$

where the subscripts indicate the x and y components of the moving points α' and

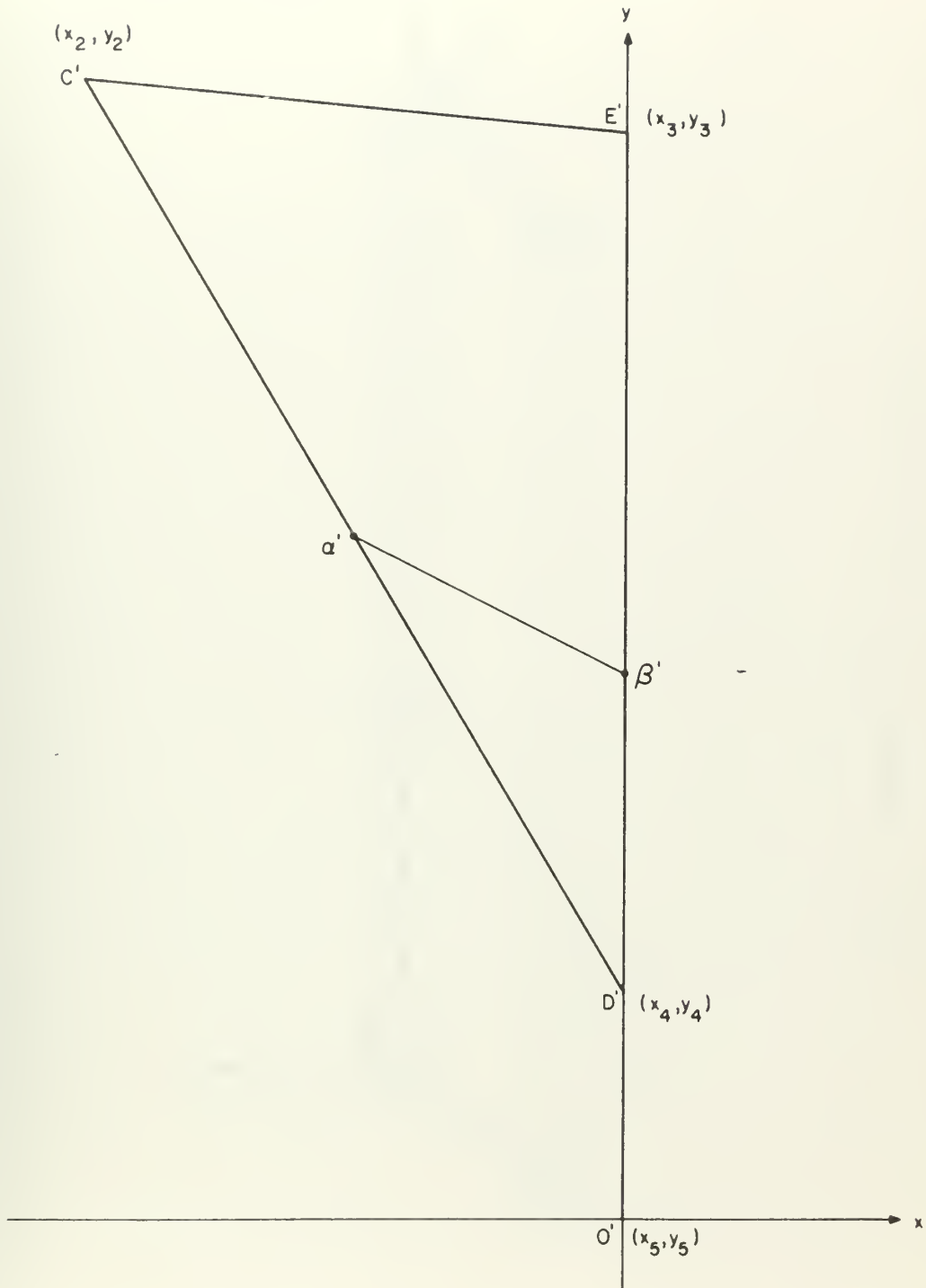


FIG. B-1 : Typical Map from (T_L, T_I, T_N) Space

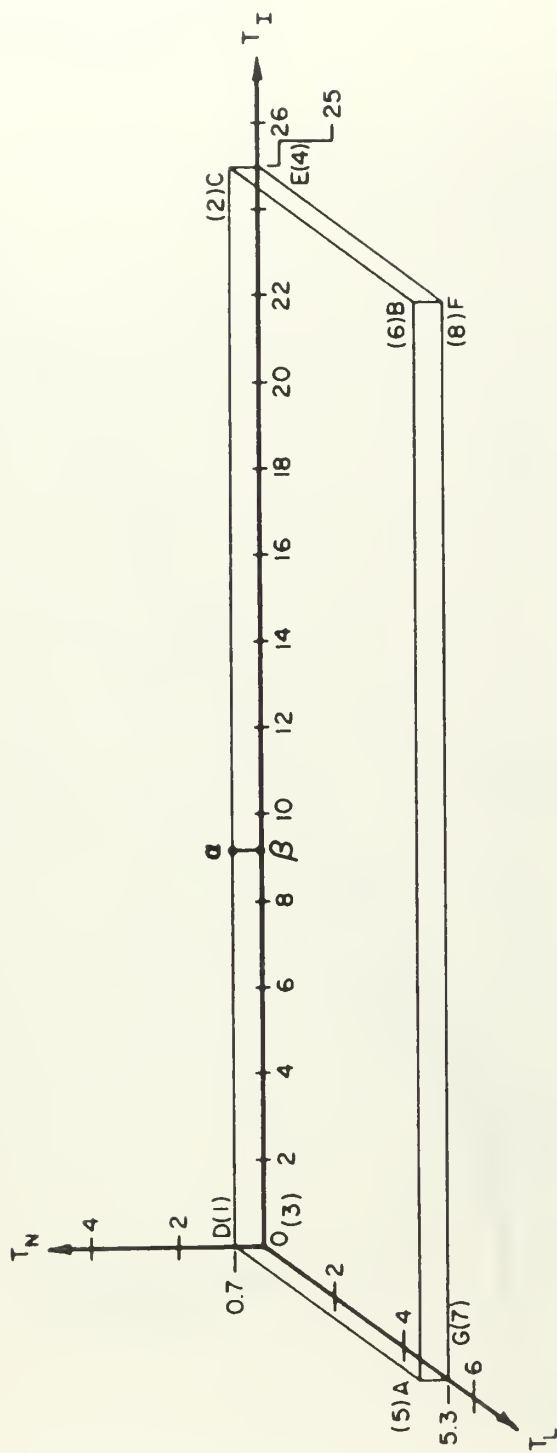


Fig. B-2 : Scanning a Face in (T_L, T_I, T_N) Space

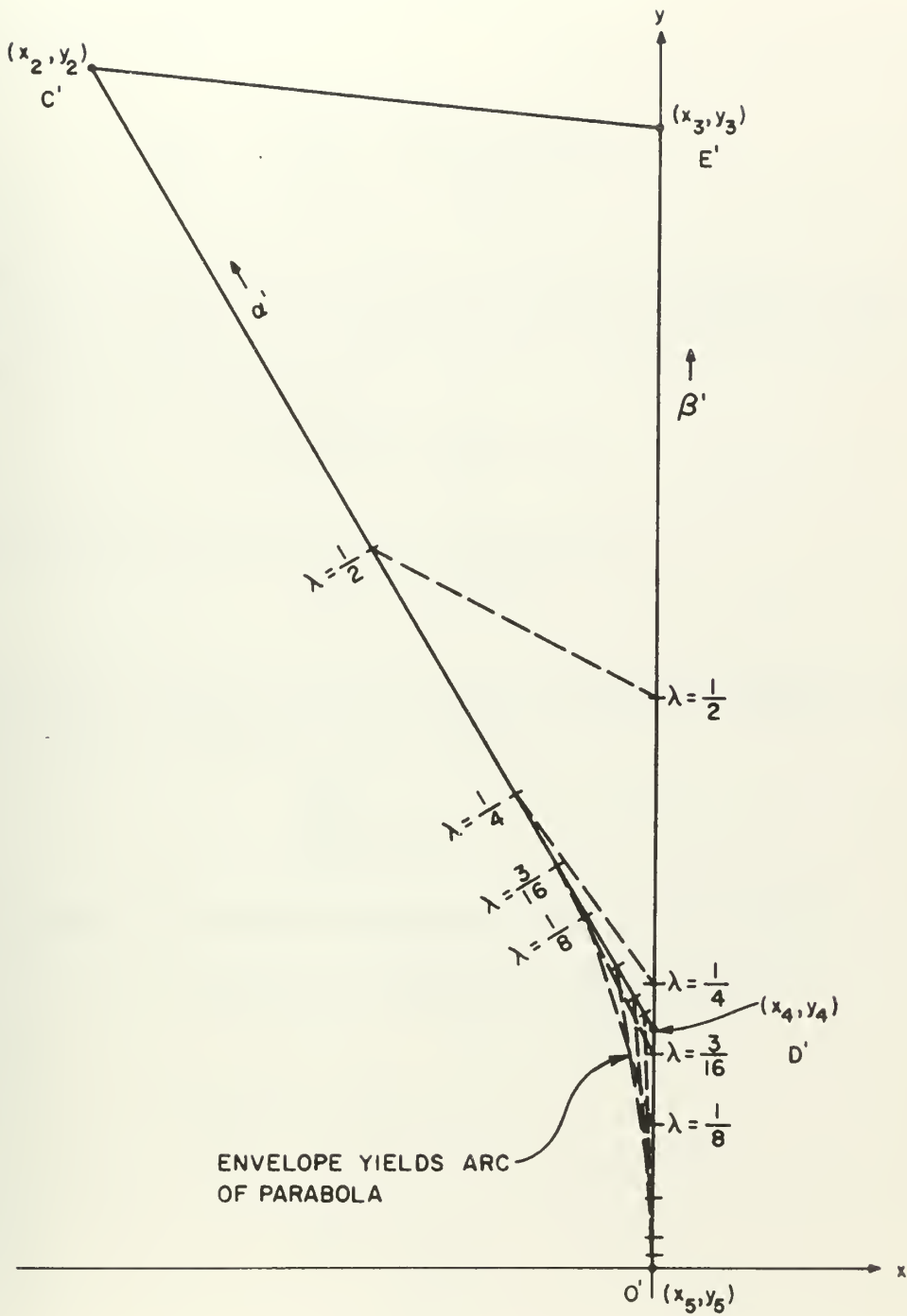


Fig. B-3 : The Parabolic Effect

β' in the space of the mapping. From these designations may be written the slope of the moving line:

$$\frac{dy}{dx} = m = \frac{\beta'_y - \alpha'_y}{\beta'_x - \alpha'_x} = \frac{\lambda y_3 - \lambda(y_2 - y_4) - y_4}{0 - \lambda x_2} \quad (\text{B-1})$$

Either by integration and evaluation of the constant of integration (when $x=0$, $y=\lambda y_3$) or through use of the slope intercept form $y=mx+b$ the equation for the family of lines may be written

$$y = \frac{y_2 - y_3 - y_4}{x_2} x + \frac{y_4}{\lambda x_2} x + \lambda y_3 \quad (\text{B-2})$$

From the theory of envelopes it is known that if⁽¹⁾

$$F(x, y, \lambda) = 0 \quad (\text{B-3})$$

represents a family of curves, then differentiating with respect to λ , holding x and y constant and equating to zero, yields

$$\frac{dF}{d\lambda}(x, y, \lambda) = 0 \quad (\text{B-4})$$

The envelope is obtained by eliminating λ between (B-3) and (B-4).

$$\therefore F(x, y, \lambda) = y - \frac{y_2 - y_3 - y_4}{x_2} x - \frac{y_4}{\lambda x_2} x - \lambda y_3 = 0$$

and

$$F_{\lambda}(x, y, \lambda) = \frac{y_4}{x_2 \lambda^2} x - y_3 = 0$$

$$y_4 x = x_2 \lambda^2 y_3$$

$$\lambda = \sqrt{\frac{y_4 x}{x_2 y_3}} \quad (\text{B-5})$$

(1) Ford, L.R., "Differential Equations", McGraw Hill, New York, 1955.

If (B-4) is substituted into (B-2) one obtains,

$$y = \frac{y_2 - y_3 - y_4}{x_2} x + \frac{y_4}{x_2} \sqrt{\frac{x_2 y_3}{y_4 x}} x + \sqrt{\frac{y_4 x}{x_2 y_3}} y_3 \quad (\text{B-6})$$

Equation (B-6) upon simplification becomes

$$y = \frac{y_2 - y_3 - y_4}{x_2} x + \sqrt{\frac{y_4 y_3}{x_2}} \sqrt{x} + \sqrt{\frac{y_4 y_3}{x_2}} \sqrt{x} \quad (\text{B-7})$$

$$y = \frac{y_2 - y_3 - y_4}{x_2} x + 2 \sqrt{\frac{y_4 y_3}{x_2}} \sqrt{x} \quad (\text{B-8})$$

A change in notation yields

$$a = \frac{y_2 - y_3 - y_4}{x_2} \quad \text{and} \quad b = 2 \sqrt{\frac{y_4 y_3}{x_2}} \quad (\text{B-9})$$

$$y - ax = b\sqrt{x}$$

Equation (B-9) squared becomes

$$y^2 - 2axy + a^2 x^2 = bx \quad (\text{B-10})$$

Equation (B-10) is a quadratic form whose parabolic nature is not obvious. Standard formulas for transformation of coordinates to principal axes form are available. ⁽¹⁾ Briefly, a quadratic form

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0 \quad (\text{B-11})$$

may have its axes rotated by an angle α so as to eliminate the cross-product

(1) Thomas, G.B., Jr., "Calculus and Analytic Geometry", Addison-Wesley, 1956.

term. A general rotation applied to the axes of Eq. (B-11) results in a new quadratic form

$$A'x'^2 + B'x'y' + C'y'^2 + D'x' + E'y' + F' = 0 \quad (\text{B-12})$$

with coefficient relations given by

$$\left. \begin{aligned} A' &= A \cos^2 \alpha + B \cos \alpha \sin \alpha + C \sin^2 \alpha \\ B' &= B(\cos^2 \alpha - \sin^2 \alpha) + 2(C-A) \sin \alpha \cos \alpha \\ C' &= A \sin^2 \alpha - B \sin \alpha \cos \alpha + C \cos^2 \alpha \\ D' &= D \cos \alpha + E \sin \alpha \\ E' &= -D \sin \alpha + E \cos \alpha \\ F' &= F \end{aligned} \right\} \quad (\text{B-13})$$

If (B-11) has coefficient $B \neq 0$ then an angle α can always be chosen such that $B' = 0$; this angle may be found from

$$\cot 2\alpha = \frac{A-C}{B} \quad (\text{B-14})$$

With the above information and some trigonometric manipulation one finds that since

$$\cot 2\alpha = \frac{1-a^2}{2a}$$

then

$$\cos 2\alpha = \frac{1-a^2}{a^2+1}$$

So,

$$\sin \alpha = \frac{a}{\sqrt{a^2+1}} \quad (\text{B-15})$$

and

$$\cos \alpha = \frac{1}{\sqrt{a^2+1}} \quad (\text{B-16})$$

Use of (B-15) and (B-16) in Eqs. (B-13) produces

$$(a^2+1)y'^2 + \frac{ab^2}{\sqrt{a^2+1}} y' - \frac{b^2}{\sqrt{a^2+1}} x' = 0 \quad (\text{B-17})$$

which has the form of a parabola. However, let

$$h = \frac{b^2}{(a^2+1)^{3/2}} \quad \text{and} \quad k = \frac{ab^2}{(a^2+1)^{3/2}}$$

then

$$y'^2 - hx' + ky' = 0$$

and completing the square

$$(y' + \frac{k}{2})^2 - \frac{b^2}{4} - hx' = 0$$

$$\therefore (y' + \frac{k}{2})^2 = h(x' + \frac{k^2}{4h})$$

Let

$$y'' = y' + \frac{k}{2}$$

and

$$x'' = x' + \frac{k^2}{4h}$$

then

$$y''^2 = hx' \quad (\text{B-18})$$

Equation (B-18) is the simple form of the equation of a parabola.

It should be noted that the area added to the end point mapping to form the complete hull is very small. In the light of this smallness, its effect upon

the gross hull over the range of T_L , and its location relative to the plots of most physical systems, there is sufficient justification to call the mapping "almost complete" in the spirit of Section 4.3.

B.2 An Exploration of T_L, T_I, T_N Space

The prospect of defining exactly the mapping of the rectangular parallelepiped of Fig. B-2 onto the Ω plane must be considered. Previous results were for the mapping of the boundaries of this volume only. As above for face ODCEO, the envelope of the region in Ω space corresponding to the mapping of every single point within the volume shown in Fig. B-2 must be obtained if both necessary and sufficient conditions for stability are to be provided for the dual Nyquist plot.

The situation $w = 1$ given in Fig. B-4 (same as Fig. 4-18) is first considered. It should be noted that face 1243 of Fig. B-2 maps into the triangle plus short line segment $1'2'4'3'$ and face 5687 maps into the triangle and short line segment $5'6'8'7'$. Indeed, as face 1243 is moved along lines 1-5, 2-6, 3-7 and 4-8 it eventually coincides with face 5687. Thus, every point of the volume in (T_L, T_I, T_N) space is scanned as face 1243 is moved along for varying T_L , i.e., as T_L goes from zero to 5.3. The effect in the complex plane plot is to have the triangular shaped mapping move from the position $1'2'4'3'$ to that of $5'6'8'7'$. This motion occurs under the constraint that point $2'$ moves along arc $2'6'$, i.e., the motion of point 2 along the line 2-6 in (T_L, T_I, T_N) space corresponds to motion of $2'$ along the circular arc $2'6'$ in the space of the mapping. Similar situations pertain to points $1', 3'$ and $4'$ in the complex plane as they move toward coincidence with $5', 7'$ and $8'$, respectively.

It is desired to show that lines $2'-4'$ and $4'-3'$, as they move toward lines $6'-8'$ and $7'-8'$, never contain points which lie outside the bounds of the hull plotted in Fig. B-4. Later, the results of the parabolic effects displayed

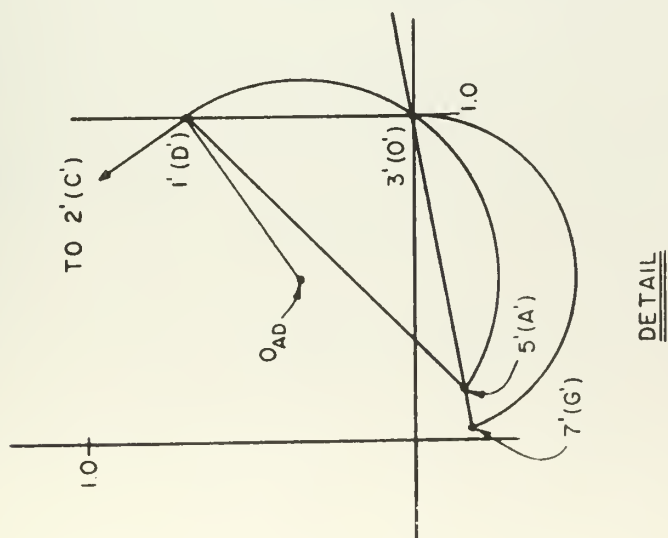
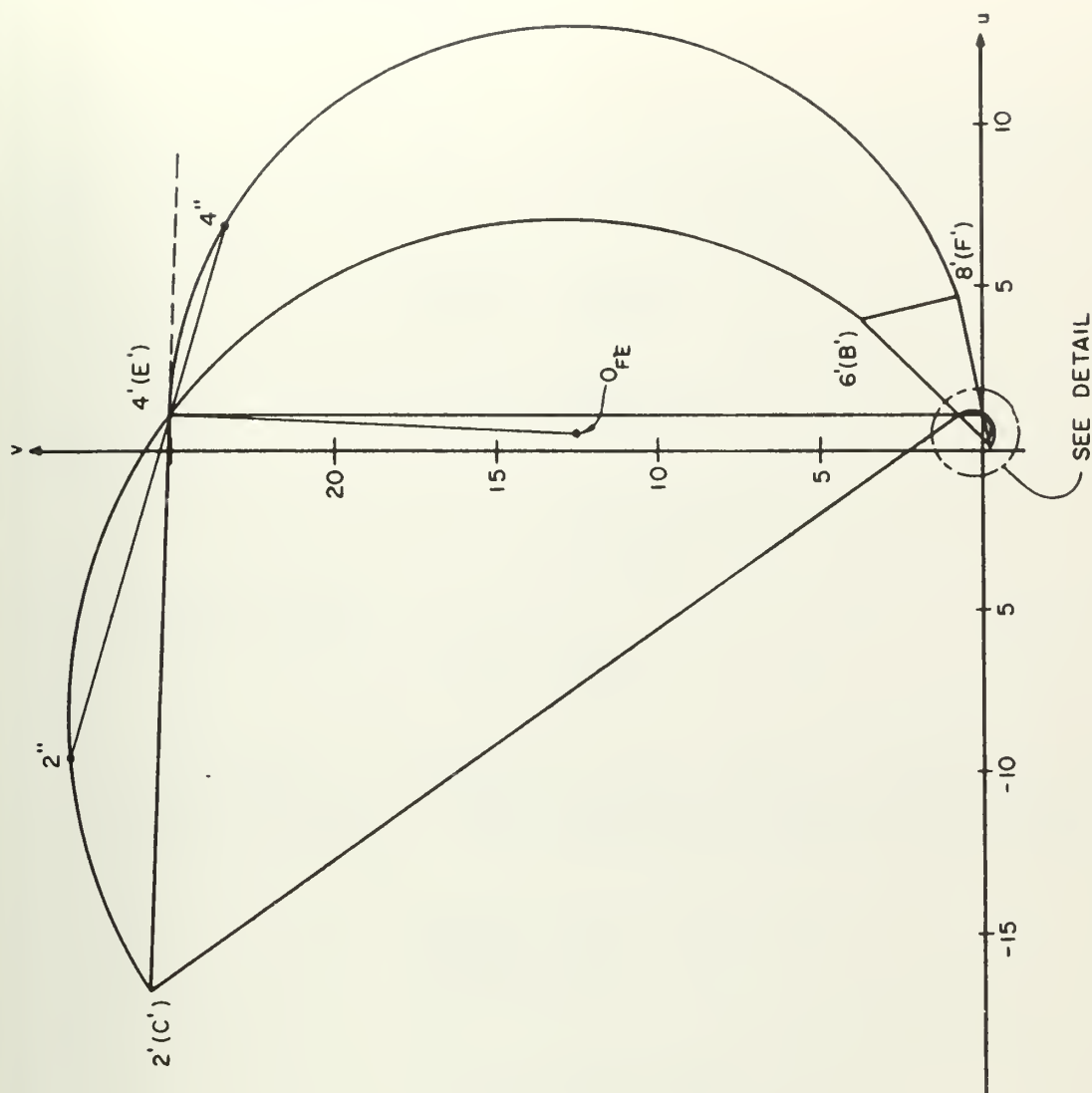


Fig. B-4 : Effect of Varying T_L

in the first part of this Appendix are considered.

For the situation of Fig. B-4, i.e., $w=1$, the line $E'C'(4'-2')$ has end points $1+j25$ and $-16.5+j25.7$, therefore its equation is

$$\frac{y-25}{25.7-25} = \frac{x-1}{-16.5-1}$$

$$y-25 = \frac{(.7)}{-17.5} (x-1)$$

or

$$y = \frac{-.7}{17.5} (x-1) + 25 \quad (B-19)$$

i.e., the line has slope

$$m_1 = \frac{-.7}{17.5}$$

The end points of the radius arm of arc $F'E'$ are $1+j25$ and the center of arc $F'E'$, i.e., $\frac{1}{2}+j12.5$. The equation of the radius arm is therefore

$$\frac{y-25}{12.5-25} = \frac{x-1}{\frac{1}{2}-1}$$

$$y-25 = \frac{-12.5}{-\frac{1}{2}} (x-1)$$

or

$$y = 25(x-1) + 25 \quad (B-20)$$

i.e., the line has slope

$$m_2 = 25$$

Since

$$m_1 m_2 = \frac{-.7}{17.5} (25) = \frac{-17.5}{17.5} = -1$$

then line $E'C'$ is perpendicular to the radius arm $0_{FE}-E'$ or $E'C'$ ($2'-4'$) is tangent to arc $F'E'$ ($4'8'$). The above is true for $w=1$. It must be shown true, in general, for any w .

For $E'C'$: The end points are $1+j25w$ and $1-17.5w^2+j25.7w$ therefore line $E'C'$ has equation

$$\frac{y-25w}{25.7w-25w} = \frac{x-1}{-17.5w^2}$$

$$\therefore m_1 = \frac{.7w}{-17.5w^2} = \frac{-.7}{17.5w}$$

For radius arm $0_{FE}-E'$ in Fig. B-4 the end points are $1+j25w$ and the center $0_{FE} = \frac{1}{2} + \frac{25w}{2}$; the equation of $0_{FE}-E'$ is then

$$\frac{y-25w}{12.5w-25w} = \frac{x-1}{\frac{1}{2}-1}$$

$$\therefore m_2 = \frac{-12.5w}{-\frac{1}{2}} = 25w$$

and

$$m_1 m_2 = \frac{-.7}{17.5w} (25w) = \frac{-17.5w}{17.5w} = -1$$

indicating, again, that $0_{FE}-E'$ is perpendicular to line $E'C'$, and therefore line $E'C'$ is tangent to arc $F'E'$.

A similar analysis answers the question, "Is $D'C'$ tangent to arc $A'D'$?"

For $D'C'$: The end points are $1+j.7w$ and $1-17.5w^2+25.7jw$ therefore the line $D'C'$ has equation

$$\frac{y-.7w}{25.7w-.7w} = \frac{x-1}{-17.5w^2}$$

$$\frac{y - .7\omega}{25\omega} = \frac{x - 1}{-17.5\omega^2}$$

$$\text{and } m_1 = \frac{25\omega}{-17.5\omega^2} \quad (\text{B-21})$$

For the radius arm of arc $A'D'$, $0_{AD-D'}$, i.e., $0_{AD-1'}$ of Fig. B-4 the end points are $1 + j.7\omega$ and the center $0_{AD} = \frac{1}{2} + j\frac{0.7}{2}\omega$ giving an equation for the radius arm of

$$\frac{y - .7\omega}{\frac{.7\omega}{2} - .7\omega} = \frac{x - 1}{\frac{1}{2} - 1}$$

$$\therefore m_2 = \frac{-.35\omega}{-\frac{1}{2}} = .7\omega \quad (\text{B-22})$$

Equations (B-21) and (B-22) give

$$m_1 m_2 = \frac{25\omega}{-17.5\omega^2} (.7\omega) = -1$$

and line $D'C'$ is tangent to arc $A'D'$.

Thus, it has been proved that initially line $2' - 4'$ is tangent to the circular arc $F'E'(4'8')$ at point $4'$. If in (T_L, T_I, T_N) space, face 1243 moves some distance along the T_L axis toward face 5687, then in the space of the mapping, point $2'$ experiences some corresponding angular excursion along the arc $2'6'$ toward $6'$. Point $4'$ experiences the same angular excursion along the arc $4'8'$ toward $8'$, etc.

Actually, it is sufficient to show that for a specified increment along T_L , equal angular increments occur on all the arc segments, i.e., it shall be proved that $T_L = x T_\ell$, $0 \leq x \leq 1$ ($T_\ell = \max T_L$) corresponds, as x varies, to arc segment increments having equal angular measure. The proof is facilitated by the passing of all circles through the origin. Thus, it is only necessary to show that a particular increase along T_L produces equivalent phase change on all the cir-

cles.

For arc $B'C'$ ($2'6'$), since

$$\Omega = \frac{1 - 17.5\omega^2 + 25.7j\omega}{1 + T_L j\omega}$$

then, for $T_L = 0$

$$\Omega = (1 - 17.5\omega^2) + j25.7\omega$$

and

$$\varphi = \tan^{-1} \frac{25.7\omega}{1 - 17.5\omega^2}$$

For $T_L = x T_\ell$

$$\Omega = \frac{(1 - 17.5\omega^2) + 25.7j\omega}{1 + x T_\ell j\omega}$$

and

$$\varphi = \tan^{-1} \frac{25.7\omega}{1 - 17.5\omega^2} - \tan^{-1} (x T_\ell \omega)$$

Therefore, the extent of arc $2'6'$ is given by

$$\Delta\varphi = \tan^{-1} x T_\ell \omega \quad (B-23)$$

For arc $F'E'$ ($4'8'$), since

$$\Omega = \frac{1 + 25j\omega}{1 + T_L j\omega}$$

then, for $T_L = 0$

$$\Omega = 1 + j25\omega$$

and

$$\varphi = \tan^{-1} 25\omega$$

For $T_L = x T_\ell$

$$\Omega = \frac{1 + j25\omega}{1 + x T_\ell j\omega}$$

and

$$\varphi = \tan^{-1} 25\omega - \tan^{-1} x T_\ell \omega$$

Therefore, the extent of arc $4'8'$ is given by

$$\Delta\varphi = \tan^{-1} x T_L \omega \quad (\text{B-24})$$

and similarly, along arc $G'O'(3'7')$ and arc $A'D'(1'5')$,

Since (B-23) equals (B-24) the statement is proved that corresponding angular motion occurs on all arc segments. Indeed, the fact follows from the properties of the bilinear transformation.⁽¹⁾

It can be mentioned at this point, also, that if T_I is set equal to $x T_i$, where $T_i = \max T_I$ and $0 \leq x \leq 1$, similar to the above exposition, then the effect again is to move correspondingly along the mapped line. For example, Fig. B-5 shows a plot, $\omega = 1$, of the situation where $0 \leq T_I \leq 12.5$, rather than $0 \leq T_I \leq 25$, as originally. The equations of lines $D'-C'$, $O'-E'$, $A'-B'$ and $G'-F'$ are the same as those in Fig. B-4, but the line segments cover only half of their previous extent. This property follows directly from the mapping

$$\Omega = \frac{(1 + j\omega T_I)(1 + j\omega T_N)}{(1 + j\omega T_L)}$$

For $\omega = \omega_1$, $T_N = \text{constant}$, $T_L = \text{constant}$

$$\begin{aligned} \Omega &= (1 + j\omega_1 T_I)(C_1 + jC_2) \\ &= (C_1 - C_2\omega_1 T_I) + j(C_2 + C_1\omega_1 T_I) \\ &= [-C_2\omega_1 + jC_1\omega_1]T_I + C_1 + jC_2 \\ &= K_1 T_I + K_2 \end{aligned} \quad (\text{B-25})$$

where $K_i = \text{some complex constant}$. Equation (B-25) demonstrates that, irrespective of the value of T_I , the mapping lies along the same straight line.

To be considered now is the case where movement has occurred corresponding to 30° of arc along $2'6'$, shown as point $2''$, and 30° of arc along $4'8'$,

(1), Churchill, R. V., "Complex Variables and Applications", 2nd Ed., McGraw-Hill, New York, 1960.

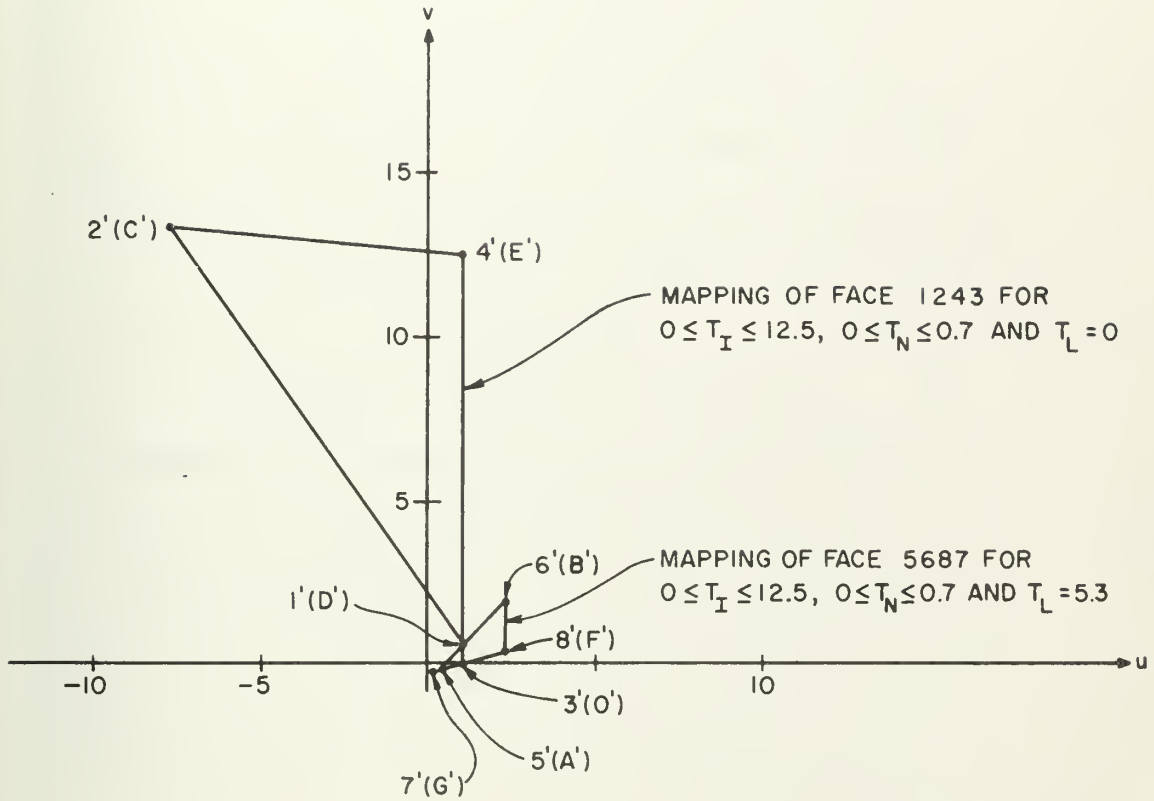


Fig. B-5 : Effect of Reduced Range on T_I

shown as $4''$, Fig. B-4. The line joining these new points corresponds in (T_L, T_I, T_N) space to an element of face 2486, Fig. B-2, which is parallel to line 2-4, or equally well, line 6-8. In the complex plane the angle $2''4'2'$ equals 15° (an inscribed angle equals one-half its intercepted arc). Also, the angle formed by the extended line $2'-4'$ tangent to arc $4'8'$ and the line $4'-4''$ equals 15° (the angle between a tangent and a chord equals one-half its intercepted arc). These two angles are therefore "vertical angles" and points $2'', 4', 4''$ can be connected by a straight line. Thus, in the space of the mapping, as $2'-4'$ moves toward $6'-8'$, the line connecting the moving points $2''-4''$ always passes through the original position of point $4'$; thus, the line NEVER passes outside the "end point" hull. Of course, this state of affairs pertains only until point $2''$ reaches $4'$, but after that it is obvious that the line $2''-4''$ is inside the hull as in the case when $2''-4''$ has finally become $6'-8'$. Similar conditions prevail in the progression of line $3'-4'$ to coincidence with $7'-8'$, i. e., the line connecting any points $3''-4''$ always passes through the original point $3'$. Thus, except for the parabolic envelope effect mentioned at the beginning of this Appendix the end point mapping of the volume in (T_L, T_I, T_N) space would be a complete one. It is granted that the area contributed by the parabolic fillet adds little to the complete hull that was not already in the end point hull. This can be seen by envisioning the triangular shaped region $1'2'4'3'$ sweeping down to $5'6'8'7'$, as T_L ranges from zero to 5.3, only keeping in mind during the sweep the presence of the parabolic fillet. If at any time the stability of a physical system is dependent upon points in the affected region they then can be taken into account. It is seen in Chapter Five that in those systems considered in this study the additional area may be disregarded. Due to the presence of this small effect the mapping is classified "almost-complete."

APPENDIX C

C.1 A Two Parameter System

The use of the three-parameter Variable-Parameter Mapping of Chapter Four for a two varying-parameter problem is considered in this Appendix. An interesting stability demonstration occurs in the following system; also, the example serves a useful purpose in the performance considerations of Chapter Six. The situation arises in the quest for a best-overall compensator which is described in detail in Chapter Six; only that portion of the design task of tutorial pertinence is investigated.

The system and compensator are as shown in Fig. C-1. The system is analyzed for values of $\zeta = 0.1$, $\omega_n = 1.57$. As in Chapter Four stability obtains if

$$1 + F(s) = 0 \quad (C-1)$$

has no right half plane zeros, where

$$F(s) = \left[\frac{s\alpha + 1}{s\beta + 1} \right] \left[\frac{V_n^2}{s(s + 2\zeta\omega_n)} \right] \quad (C-2)$$

If (C-1) is multiplied by

$$\frac{s\beta + 1}{s\alpha + 1} = \frac{1}{V} \quad (C-3)$$

the fixed and varying portions for the plots are obtained. That is,

$$\frac{s\beta + 1}{s\alpha + 1} + \frac{V_n^2}{s(s + 2\zeta\omega_n)} = 0 \quad (C-4)$$

and

$$\frac{V_n}{V} = \frac{s\beta + 1}{s\alpha + 1} \quad (C-5)$$

where V_n is again taken to be unity as in Section 4.2, i. e., $\alpha = \beta = 0$.

Also,

$$L_n = \frac{V_n^2}{s(s + 2\zeta\omega_n)} \quad (C-6)$$

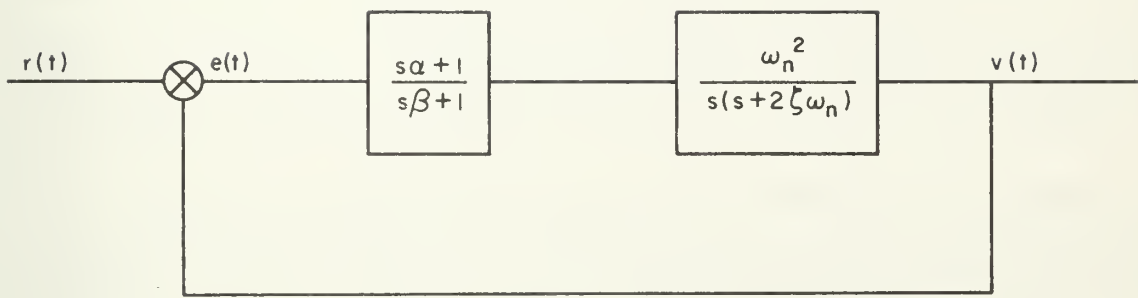


Fig. C-1 : Second-Order System with Compensator

This results in the standard plotting structure

$$\frac{V}{V_n} = -L_n \quad (C-7)$$

C.2 The Degenerate Mapping

The constraints on permissible α and β are taken to be $0 \leq \alpha \leq 5.3$ and $0 \leq \beta \leq 25$. It is apparent that the $\frac{V}{V_n}$ mapping is obtained from the image of face 3487 shown in Fig. 4-16. This mapping is actually shown in Fig. 4-20, for $w=1.0$, i.e., it is represented by the bounds formed by line 3'-4', arc 4'-8', line 8'-7' and arc 7'-3'. The degenerate mapping just outlined is shown in Fig. C-2 and again in Fig. C-3, along with the plot of $-L_n$ given by the negative of (C-6).

The primary purpose of this example is to demonstrate that sufficient caution must be exercised in interpreting the plots to preclude succumbing to the temptation of making stability decisions based upon limited frequency information. One must guard against using the thumb rule of Section 4.2 superficially, by considering only one or two frequency points. The remainder of the example clarifies this notion. Upon observation of Fig. C-3 one might infer that some set of α, β , such as $\alpha=2, \beta=25$, would result in a stable system. That is, for $\beta=25$, as α varies, the image point of $\frac{V}{V_n}$ runs along arc segment 4'-8', and for $\alpha=2$ if the image point is "to the left" of $-L_n$ (tracing $-L_n$ as frequency increases), such a condition seems to imply a stable situation. The point corresponding to $\alpha=2, \beta=25$ and, of course, $w=1.0$ is shown as $M(j1)$ in Fig. C-3 where

$$M = \frac{j\omega(25) + 1}{j\omega(2) + 1} \quad (C-8)$$

That is,

$$M = \frac{V}{V_n} (\alpha=2, \beta=25)$$

and it is observed that $M(j1)$ does fall to the left of $-L_n$ implying stability.

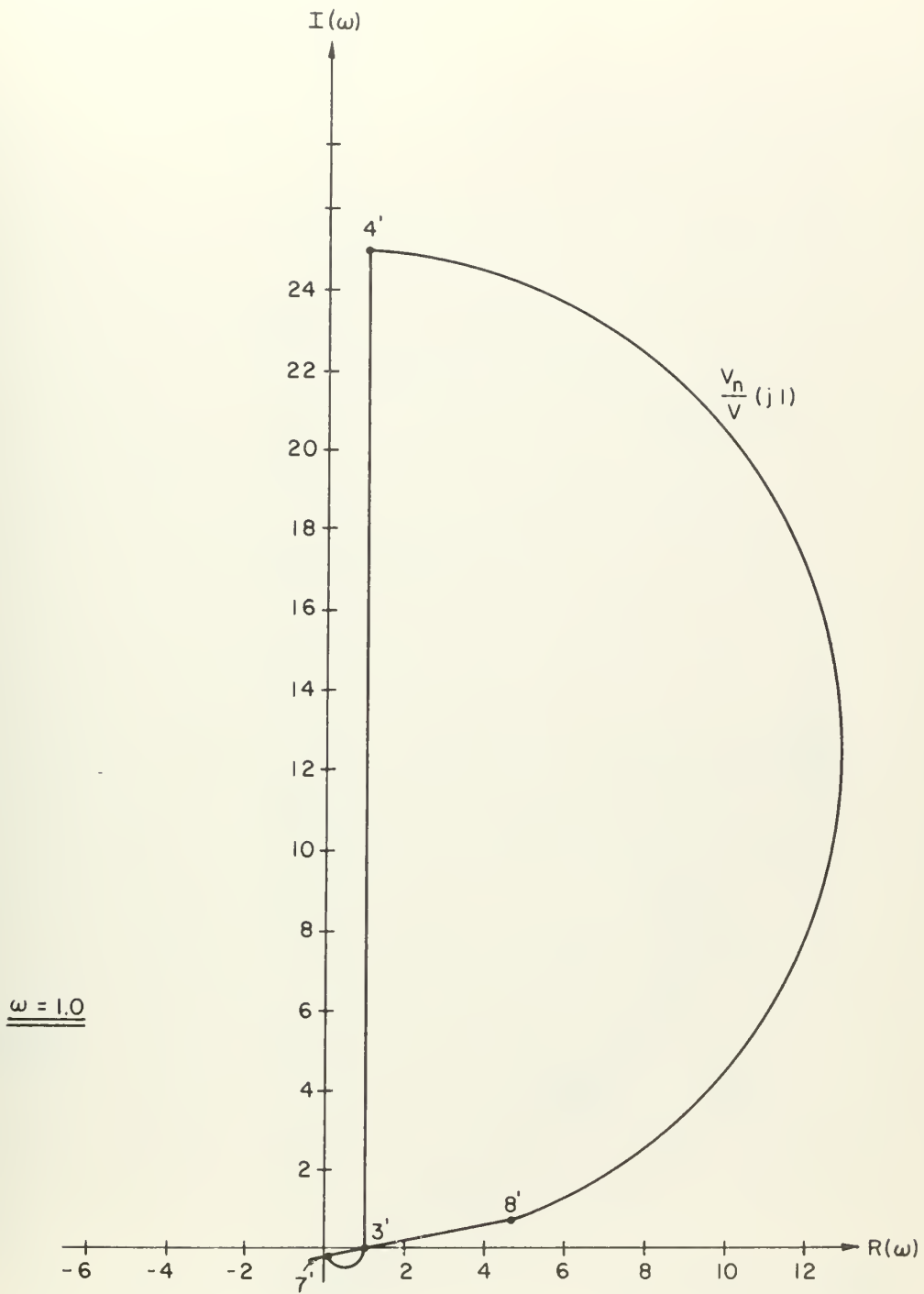


Fig. C-2 : The Degenerate Three-Parameter Mapping
for Two-Parameter Analysis

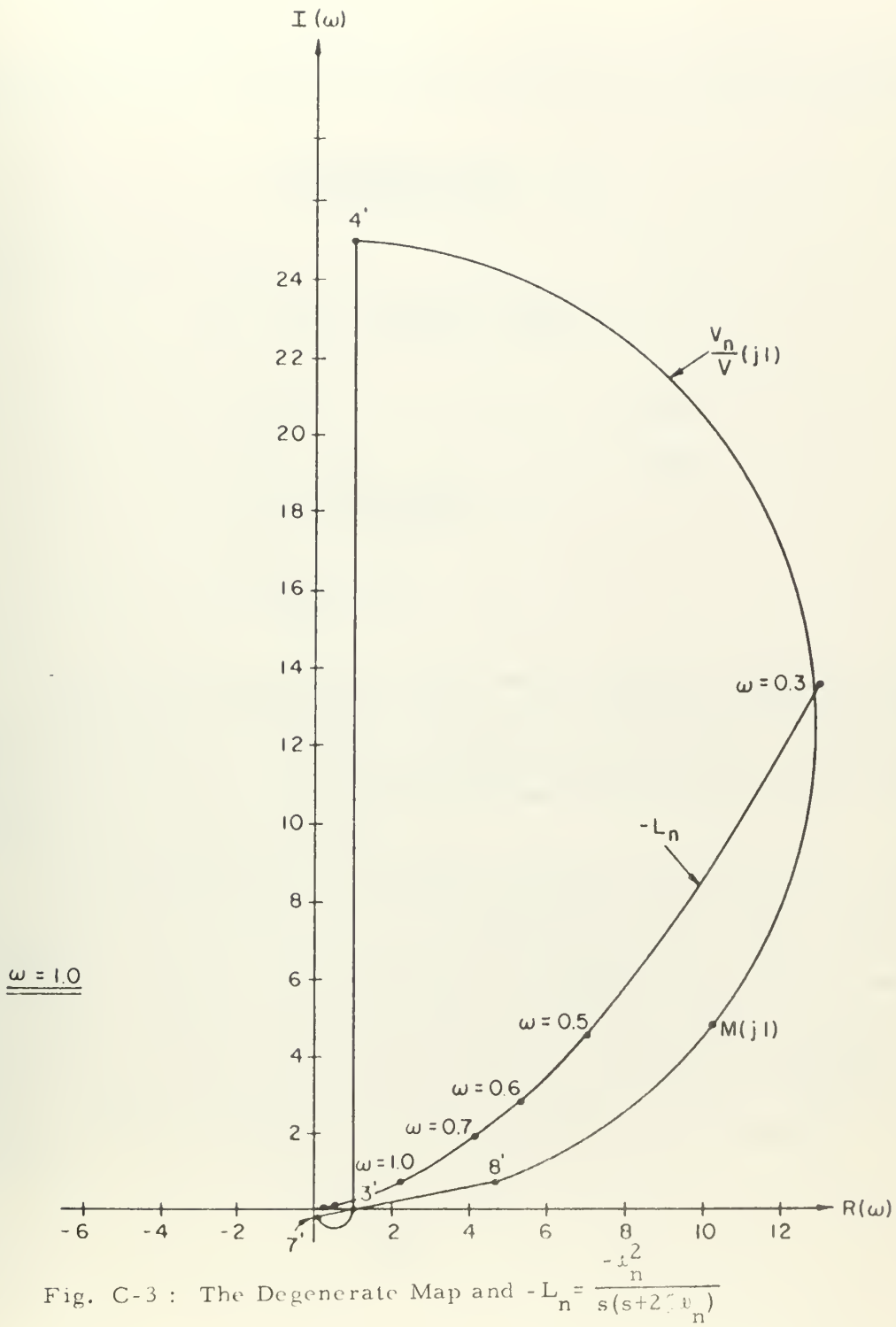


Fig. C-3 : The Degenerate Map and $-L_n = \frac{n}{s(s+2v_n)}$

$$\tau = 0.1, x_n = 1.57$$

An actual stability check for these specified circumstances can be obtained from the Routh test. From (C-1) and the characteristics given one obtains

$$1 + \frac{(2.46)(s^2 + 1)}{s(s^2 + 1)(s + .314)} = 0 \quad (\text{C-10})$$

From (C-10) there follows

$$s^3 + .354s^2 + .20956s + .0985 = 0$$

| | | |
|-------|--|--------|
| s^3 | 1 | .20956 |
| s^2 | .354 | .0985 |
| s^1 | $\frac{.0742 - .0985}{.354} \rightarrow (-)$ | |
| s^0 | $\frac{(-)}{(-)} \rightarrow (+)$ | |

The two sign changes in the array above signify two right-half-plane roots in the characteristic equation and, consequently, instability. The apparent inconsistency between the graphical indications and the analytical results can be cleared away by a more detailed plotting of the $M(j\omega)$ locus at more frequencies, particularly those less than unity.

The pertinent points of the locus for $M(j\omega)$ are shown in Fig. C-4 along with $-L_n(j\omega)$. It is noted that the "difference vector" drawn from $-L_n$ to M undergoes a clockwise rotation between $M(0.3)$ and $M(0.7)$ thus indicating instability and consistency with the Routh test. The difference vector actually makes two complete clockwise rotations corresponding to the existence of the two right-half-plane roots in the characteristic equation. To observe the complete revolutions the negative frequency branches of $-L_n$ and M must be considered; also, the effects of $s=0$ are handled by the standard technique of setting

$$s = \lim_{R \rightarrow 0} R e^{j\theta}$$

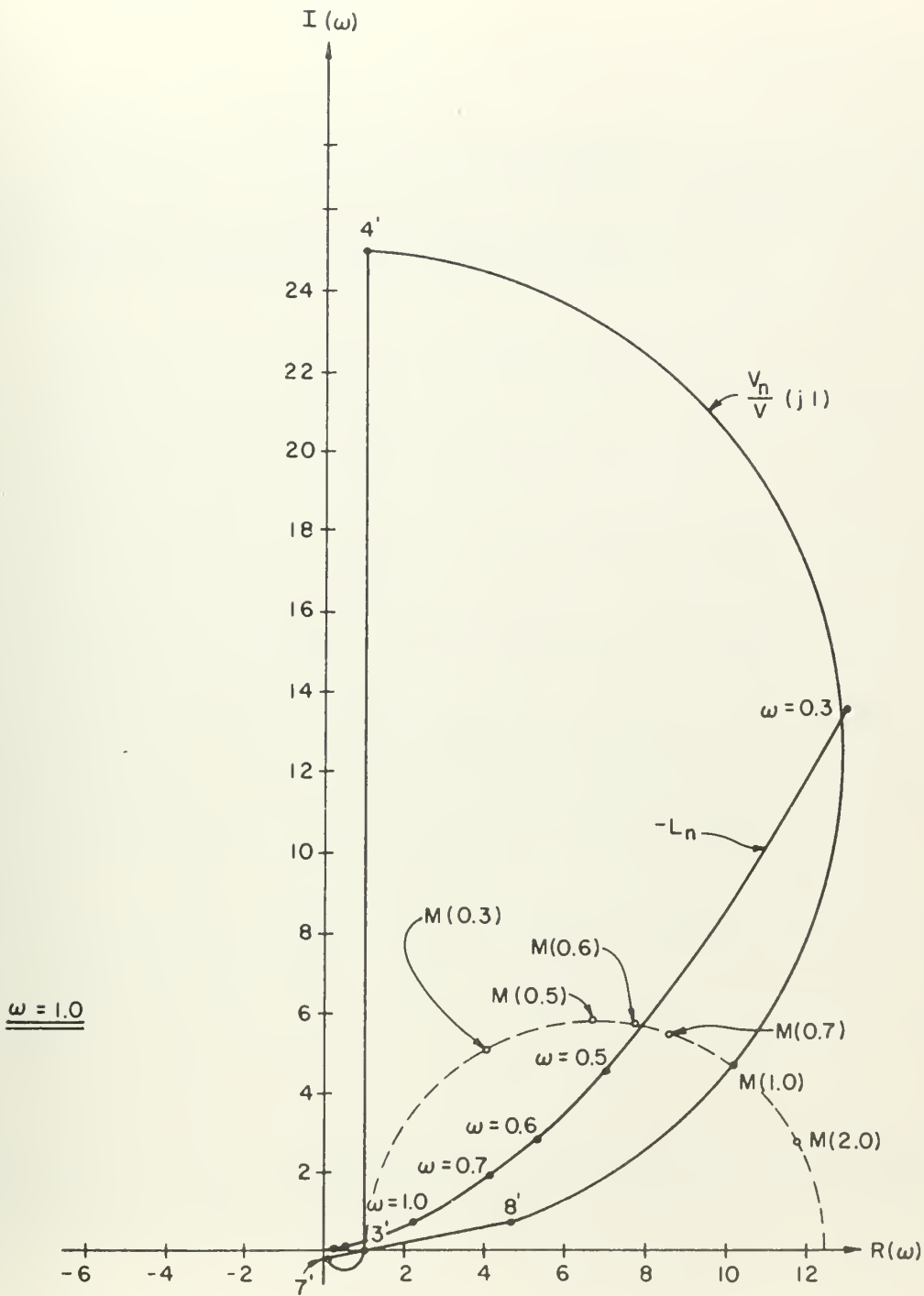


Fig. C-4 : Two-Parameter Variation for $\alpha=2$, $\beta=25$; also shows $-L_n$

where θ varies from $-\frac{\pi}{2}$ to $\frac{\pi}{2}$ and the origin of the s -plane is circumvented by a small semi-circle of radius $R^{(1)}$.

Difficulties with "encirclement" interpretation can normally be resolved by sketching the entire locus of $-L_n$ for positive and negative frequencies (similarly for the varying portion); the excursions of the difference vector between these two loci may then be more readily observed.

⁽¹⁾ Del Toro, V. and S. R. Parker, Principles of Control Systems Engineering, 329-330, McGraw-Hill, N. Y., 1960.

BIBLIOGRAPHY

1. Arthur, G. R., A Note on the Approach of Narrow Band Noise after a Non-Linear Device to a Normal Probability Density, J. Appl. Phys., Vol. 23, 1952.
2. Ashkenas, I.L. and D. T. McRuer, A Theory of Handling Qualities Derived from Pilot-Vehicle System Considerations, Aerospace Engineering, Feb., 1962.
3. Bahrick, H. P., P.M. Fitts and G.E. Briggs, Learning Curves - Facts or Artifacts, Psych. Bull., 54, 256-268, 1957.
4. Bekey, G. A., An Investigation of Sampled Data Models of the Human Operator in a Control System, (Ph. D. Thesis, UCLA, Jan., 1963) also issued as ASD-TDR-62-36, WADC, May, 1962.
5. Bekey, G. A., H. F. Meissinger and R. E. Rose, A Study of Model Matching Techniques for the Determination of Parameters In Human Pilot Models, NASA CR-143, Jan., 1965.
6. Bellman, R., "Adaptive Control Processes - A Guided Tour", Princeton University Press, Princeton, New Jersey, 1961.
7. Belsley, S. E., Man-Machine System Simulation for Flight Vehicles, IEEE Trans. on Human Factors in Electronics, Sept., 1963.
8. Bergeron, H. P., J. K. Kincaid and J. J. Adams, Measured Human Transfer Functions in Simulated Single-Degree-of-Freedom Nonlinear Control Systems, NASA TN D-2569, Jan., 1965.
9. Birmingham, H. P. and F. V. Taylor, A Human Engineering Approach to the Design of Man-Operated Continuous Control Systems, NRL Rep. 5333, Apr., 1954, (also published as A Design Philosophy for Man-Machine Control Systems, Proc. IRE, Dec., 1954).
10. Box, G. E. P. and K. B. Wilson, The Experimental Attainment of Optimal Conditions, J. Roy. Stat. Soc., B13, 1 (1951).
11. Blecher, F. H., Design Principles for Multiple-Loop Transistor Feedback Amplifiers, Bell Labs Memo. 57-2172-6, Nov., 1957; also appears in Proceedings of the National Electronics Conference, Vol. XIII, Oct., 1959. (The stability criterion in this reference is due originally to M. L. Curtis).
12. Booton, R. C., The Analysis of Non-Linear Control Systems with Random Inputs, Proc. Symp. on Non-Linear Circuit Analysis, Polytechnic Institute of Brooklyn, N. Y., Apr., 1953.

13. Burrows, A.A., Editor, Complex Vehicular Controls, Proc. of a Symposium sponsored by ONR and held at Farnborough, England, May, 1962.
14. Choksy, N.H., Time Lag Systems, "Progress in Control Engineering", Vol. 1, Academic Press, N. Y., 1962.
15. Churchill, R.V., "Complex Variables and Applications", 2nd Ed., McGraw-Hill, New York, 1960.
16. Cole, G. et al, Study of Pilot-Controller Integration for Emergency Conditions, TDR-63-4092, WPAFB, Ohio, 1963.
17. Cooper, G.E., Understanding and Interpreting Pilot Opinion, Aero. Eng. Rev., Vol. 16, No. 3, Mar., 1957.
18. Cornell Aeronautical Lab. Inc., Simulator Evaluation of Airplane Longitudinal Responses for the Instrument-Landing Approach, Tech. Documentary Rept. No. FDL-TDR-64-84, Oct., 1964.
19. Dander, V.A., Predicting Pilot Ratings of Multi-Axis Control Tasks from Single Axis Data, IEEE Trans. on Hum. Factors in Electronics, Sept., 1963.
20. Del Toro, V. and S.R. Parker, Principles of Control Systems Engineering, 329-330, McGraw-Hill, N. Y., 1960.
21. Dzung, L.H., The Stability Criterion, "Automatic and Manual Control", Tustin, A., Cranfield Conference papers, Butterworths (publishers), 1952.
22. Elkind, J.I., Tracking Response Characteristics of the Human Operator, Human Factors Operations Research Lab. Rpt. No. 40, PRP-2, Sept., 1953.
23. Elkind, J.I., Characteristics of Simple Manual Control Systems, (Ph. D. Thesis, M.I. T., Apr., 1956) also issued as M.I. T. Lincoln Lab. Report, No. 111.
24. Elkind, J.I. and D.M. Green, Measurement of Time-Varying and Non-Linear Dynamic Characteristics of Human Pilots, Aeronautical Systems Division, Tech. Rpt. No. 61-225, Dec., 1961.
25. Elkind, J.I. et al, Evaluation of a Technique for Determining Time-Invariant and Time-Variant Dynamic Characteristics of Human Pilots, NASA, TND-1897, May, 1963.
26. Elkind, J.I. and D.L. Darley, The Normality of Signals and Describing Function Measurements of Simple Manual Control Systems, IEEE Trans. on Human Factors in Electronics, Sept., 1963.

27. Ferrell, W. R., Remote Manipulation with Transmission Delay, NASA TN D-2665, Feb., 1965.
28. Ford, L. R., "Differential Equations", McGraw-Hill, New York, 1955.
29. Gardner, M. F. and J. L. Barnes, "Transients in Linear Systems", Vol. 1, Wiley, New York, 1942.
30. Gibbs, C. B., Controller Design: Interactions of Controlling Limbs, Time-Lags and Gains in Positional and Velocity Systems, Ergonomics, Vol. 5, No. 2, Apr., 1962.
31. Gibson, J. E., "Nonlinear Automatic Control", McGraw-Hill, New York, 1963.
32. Gibson, J. E. et al, Philosophy and State of the Art of Learning Control Systems, CISL, Purdue University, Nov., 1963.
33. Graham, D. and D. T. McRuer, "Analysis of Nonlinear Control Systems", John Wiley and Sons, 1961.
34. Hall, I. A. M., Effects of Controlled Element on the Human Pilot, Princeton Univ., WADC Tech. Dept. No. 57-509, Aug., 1958.
35. Hickey, A. E. and W. C. Blair, Man as a Monitor, Human Factors, Sept., 1958.
36. Hooke, R. and T. A. Jeeves, Direct Search Solution of Numerical and Statistical Problems, J. Assoc. of Comp. Mach.; 8, 2, Apr., 1961.
37. Horowitz, I. M., "Synthesis of Feedback Systems", Academic Press, N. Y., 1963.
38. Jackson, A. S., Synthesis of a Linear Quasi-Transfer Function for the Operator in Man-Machine Systems, WESCON Convention Record, pt. 4, 1958.
39. James, H. M., N. B. Nichols and R. S. Phillips, "Theory of Servomechanisms", McGraw-Hill, N. Y., 1947.
40. Johnston, D. E. and D. H. Weir, Study of Pilot-Vehicle-Controller Integration for a Minimum Complexity AFCS, Tech. Rept. No. 127-1, Systems Technology, Inc., Inglewood, Calif., July, 1964.
41. Jones, P., Stability of Feedback Systems Using Dual Nyquist Diagram, Trans. IRE Prof. Group on Circuit Theory, Mar., 1954.
42. Krendel, E. S. and D. T. McRuer, A Servomechanisms Approach to Skill Development, J. Franklin Institute, Vol. 269, No. 1, Jan., 1960.

43. Marienfeld, Horst, "Simulationstechnik-Allgemeine Grundlagen und Anwendung in der Starrflüglertechnik", Verlag des Vereins Deutscher Ingenieure, 1964.
44. McRuer, D. T. and E. S. Krendel, Dynamic Response of Human Operators, WADC Tech. Rpt. 56-524, Oct., 1957.
45. McRuer, D. T. and E. S. Krendel, The Human Operator as a Servo System Element, J. Franklin Institute, Vol. 267, Nos. 5 and 6, May and June, 1959.
46. McRuer, D. T., I. L. Ashkenas and C. L. Guerre, A Systems Analysis View of Longitudinal Flying Qualities, WADD Tech. Rept. 60-43, Wright Air Development Division, Jan., 1960.
47. McRuer, D. T., and I. L. Ashkenas, Design Implications of the Human Transfer Function, Aerospace Engineering, Sept., 1962.
48. Mitchell, M. B., Appendix to Manual Control, A Survey of Human Operator Models for Manual Control, prepared for Office of Naval Research, Wash., D. C., 1964.
49. Muckler, F. A. and R. W. Obermayer, The Human Operator, International Science and Technology, July, 1964.
50. Muckler, F. A. and R. W. Obermayer, Control System Lags and Man-Machine System Performance, NASA CR-83, July, 1964.
51. Muckler, F. A. and R. W. Obermayer, The Use of Man in Booster Guidance and Control, NASA CR-81, July, 1964.
52. Myers, B. R., A Useful Extension of the Nyquist Criterion to the Stability Analysis of Multiloop Feedback Amplifiers, Proc. 4th Midwest Symposium on Circuit Theory, Marquette University, Dec., 1959.
53. Newton, G. C., L. A. Gould and J. F. Kaiser, "Analytical Design of Linear Feedback Controls", John Wiley and Sons, New York, 1957.
54. Obermayer, R. W. and F. A. Muckler, On the Inverse Optimal Control Problem in Manual Control Systems, NASA CR-208, Apr., 1965.
55. Pennington, J. E., H. G. Hatch, Jr., E. R. Long and J. B. Cobb, Visual Aspects of a Full-Size Pilot-Controlled Simulation of the Gemini-Agena Docking, NASA TN D-2632, Feb., 1965.
56. Roig, R. W., A Comparison Between Human Operator and Optimum Linear Controller RMS-Error Performance, IRE Transactions on Human Factors in Electronics, Mar., 1962.

57. Russell, L., Characteristics of the Human as a Linear Servo Element (M. S. Thesis, M. I. T., May, 1951).
58. Sadoff, M., A Study of a Pilot's Ability to Control During Simulated Stability Augmentation System Failures, NASA, TN D-1552, Nov., 1962.
59. Satche, M., Discussion, J. Appl. Mech., p. 419, Dec., 1949.
60. Senders, J. W., The Human Operator as a Monitor and Controller of Multi-degree of Freedom Systems, IEEE Transactions on Human Factors in Electronics, Sept., 1964.
61. Sheridan, T. B., Studies of Adaptive Characteristics of the Human Controller Air Force Systems Command ESD-TDR 62-351, Dec., 1962.
62. Sheridan, T. B., The Human Operator in Control Instrumentation, "Progress in Control Engineering", Vol. 1, Academic Press, New York, 1962.
63. Sherman, S. Non-Mean-Square Error Criteria, IRE Transactions on Information Theory, Sept., 1958,
64. Smith, R. H., An Experimental Determination of the Limits of Pilot Controllability for Unstable, Oscillatory Second Order Systems, Thesis for the degree of Aeronautical Engineer, Univ. of Cincinnati, 1961 (also appeared as a paper, On the Limits of Manual Control, IEEE Trans. on Hum. Factors in Electronics, Sept., 1963.
65. Sobczyk, A., Aided Tracking, M. I. T. Radiation Laboratory Report 430, Sept., 1943.
66. Stewart, R. M., A Simple Graphical Method for Constructing Families of Nyquist Diagrams, J. Aero. Sciences, p. 767, Nov., 1951.
67. Summers, L. G. and K. Ziedman, A Study of Manual Control Methodology with Annotated Bibliography, NASA Contractor Report CR-125, Nov., 1964.
68. Taylor, F. V. and H. P. Birmingham, That Confounded System Performance Measure - A Demonstration, Psychological Review, 66, 178-182, 1959.
69. Taylor, F. V., Human Engineering and Psychology, in "Psychology: A Study of a Science", Koch. S., (Ed.), Vol. 5, McGraw-Hill, N. Y., 1963.
70. Thomas, G. B., Jr., "Calculus and Analytic Geometry", Addison-Wesley, 1956.
71. Thomas, R. E. et al, The Effect of Various Levels of Automation on Human Operators' Performance in Man-Machine Systems, WADD Tech. Rpt. 60-618, Feb., 1961.

72. Tustin, A., An Investigation of the Operator's Response in Manual Control of a Power Driven Gun, C. S. Memorandum No. 169, Metropolitan-Vickers Electrical Co. Ltd., Attercliffe Common Works, Sheffield, England, Aug., 1944.
73. USAF Instrument Pilot School, Randolph AFB, Texas, Control-Display Pilot Factors Program, Instrument Evaluation Project Report 63-1, Dec., 1963.
74. Wierwille, W. W. and G. A. Gagne, A Theory for the Optimal Deterministic Characterization of the Time-Varying Dynamics of the Human Operator, NASA CR-170, Feb., 1965.
75. Wilde, D. J., "Optimum Seeking Methods", Prentice-Hall, New Jersey, 1964.
76. Young, L. R. and L. Stark, Variable Feedback Experiments Testing a Sampled Data Model for Eye Tracking Movements, IEEE Transactions on Human Factors in Electronics, Sept., 1963.
77. Young, L. R., D. M. Green, J. I. Elkind and J. A. Kelly, Adaptive Dynamic Response Characteristics in Simple Manual Control, IEEE Transactions on Human Factors in Electronics, Sept., 1964.
78. Young, L. R. and L. Stark, Biological Control Systems - A Critical Review and Evaluation, Developments in Manual Control, NASA CR-190, Mar., 1965.
79. Wilde, D. J., "Optimum Seeking Methods", Prentice-Hall, Englewood Cliffs, N. J., 1964.
80. Wilde, R. W. and J. H. Westcott, The Characteristics of the Human Operator Engaged in a Tracking Task, Automatica, Vol. 1, No. 1, Jan. - Mar., 1963.

The text and drawings of this thesis
were duplicated by the Offset Printing process.

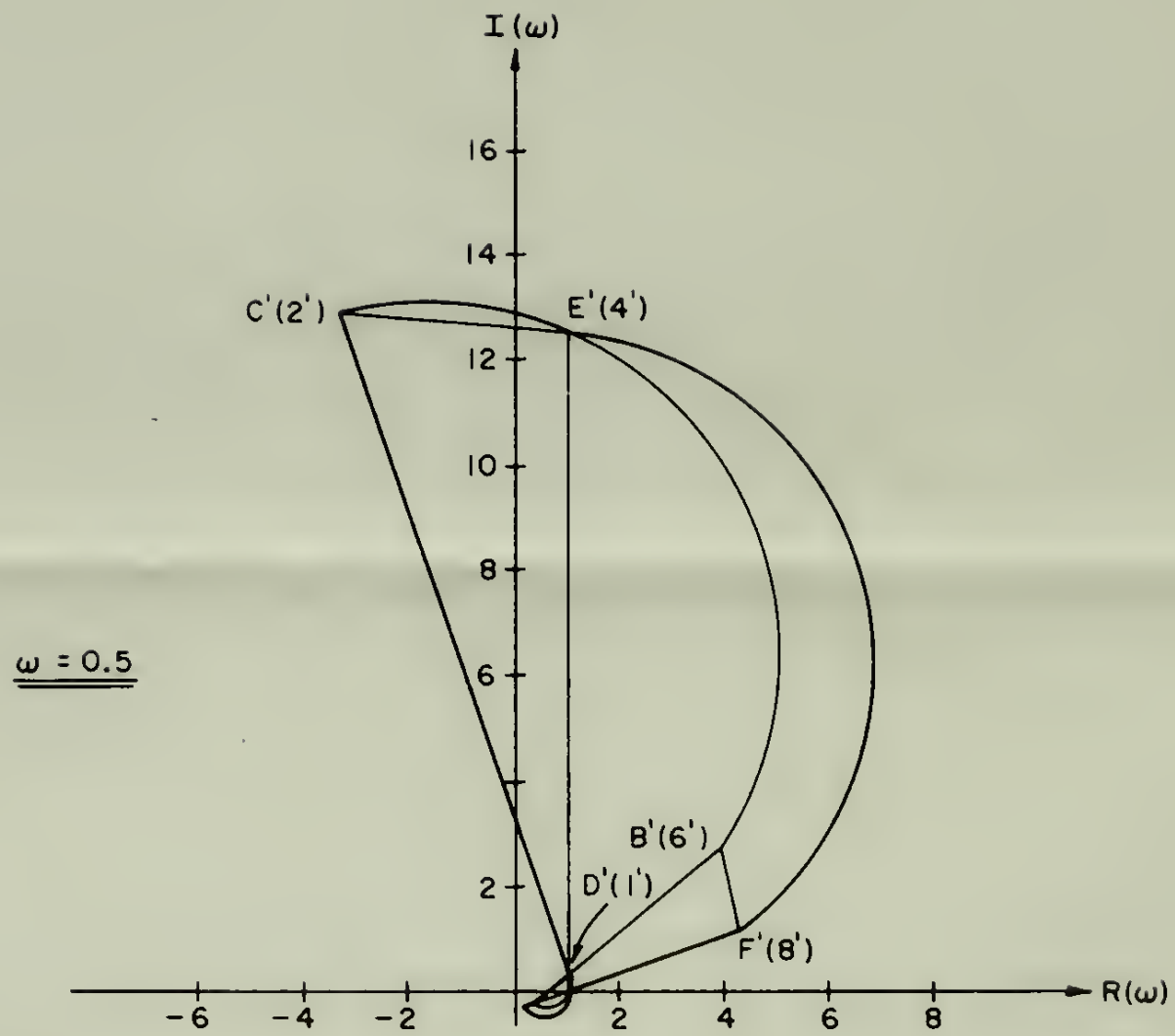


Fig. 4-19: Variable Parameter Mapping

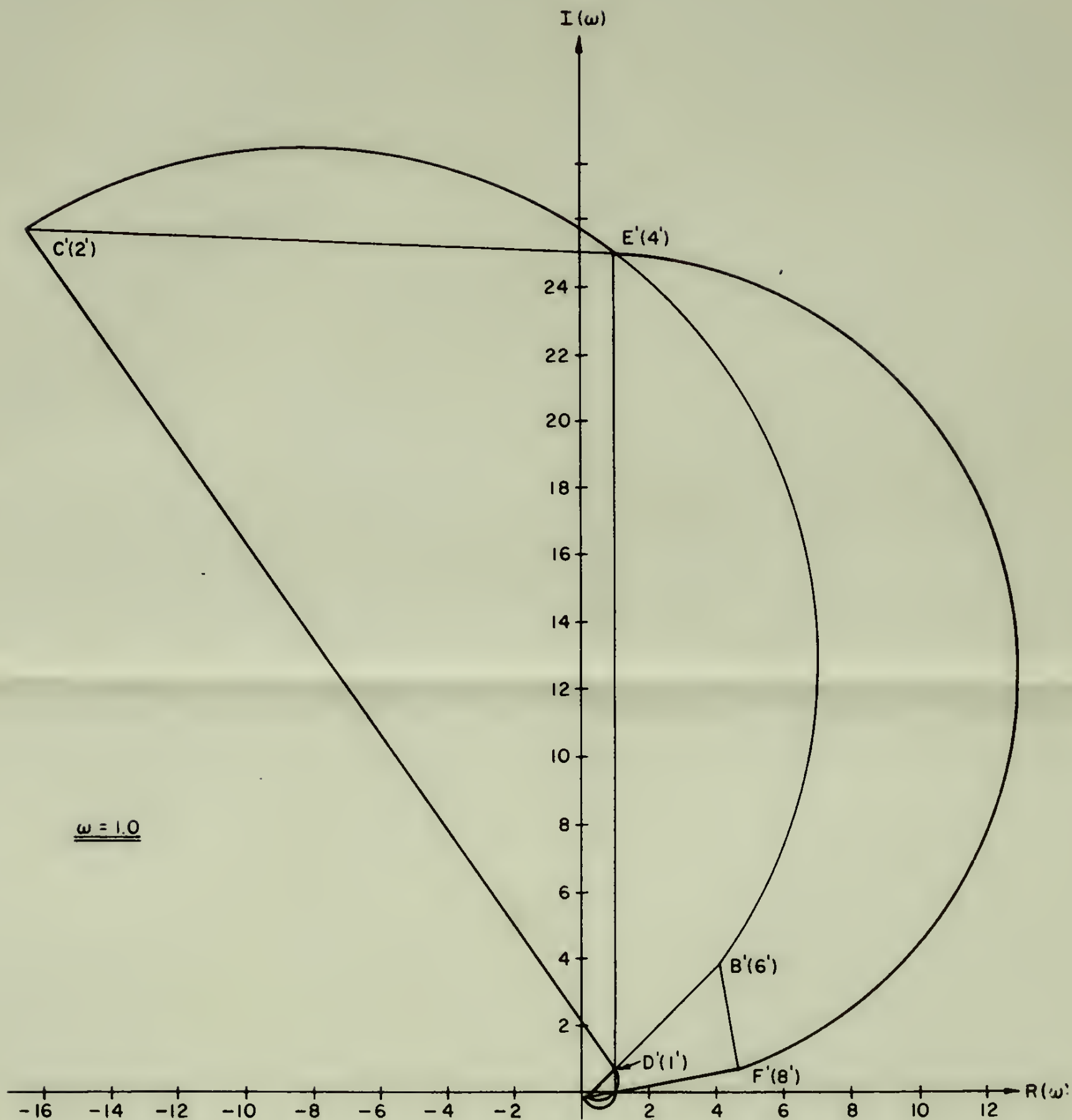


Fig.4-20 : Variable Parameter Mapping

Library
 P. S. Narayana Postgraduate School
 Mysore-57. 500.001

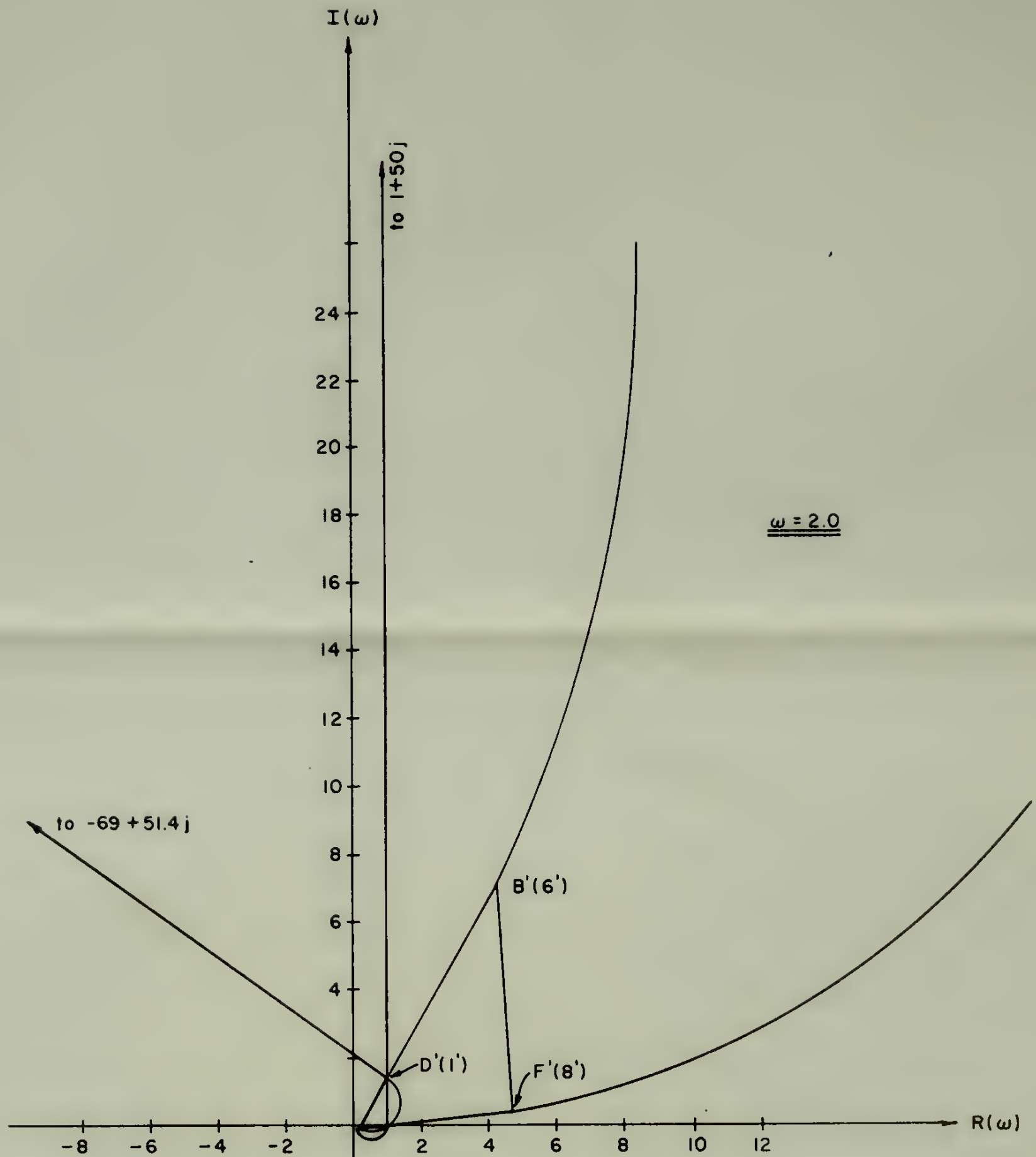


Fig. 4-21 : Variable Parameter Mapping

Library
 U. S. Naval Postgraduate School
 Monterey, California

Thesis
557

105

Library
U. S. Naval Postgraduate School
Monterey, CA

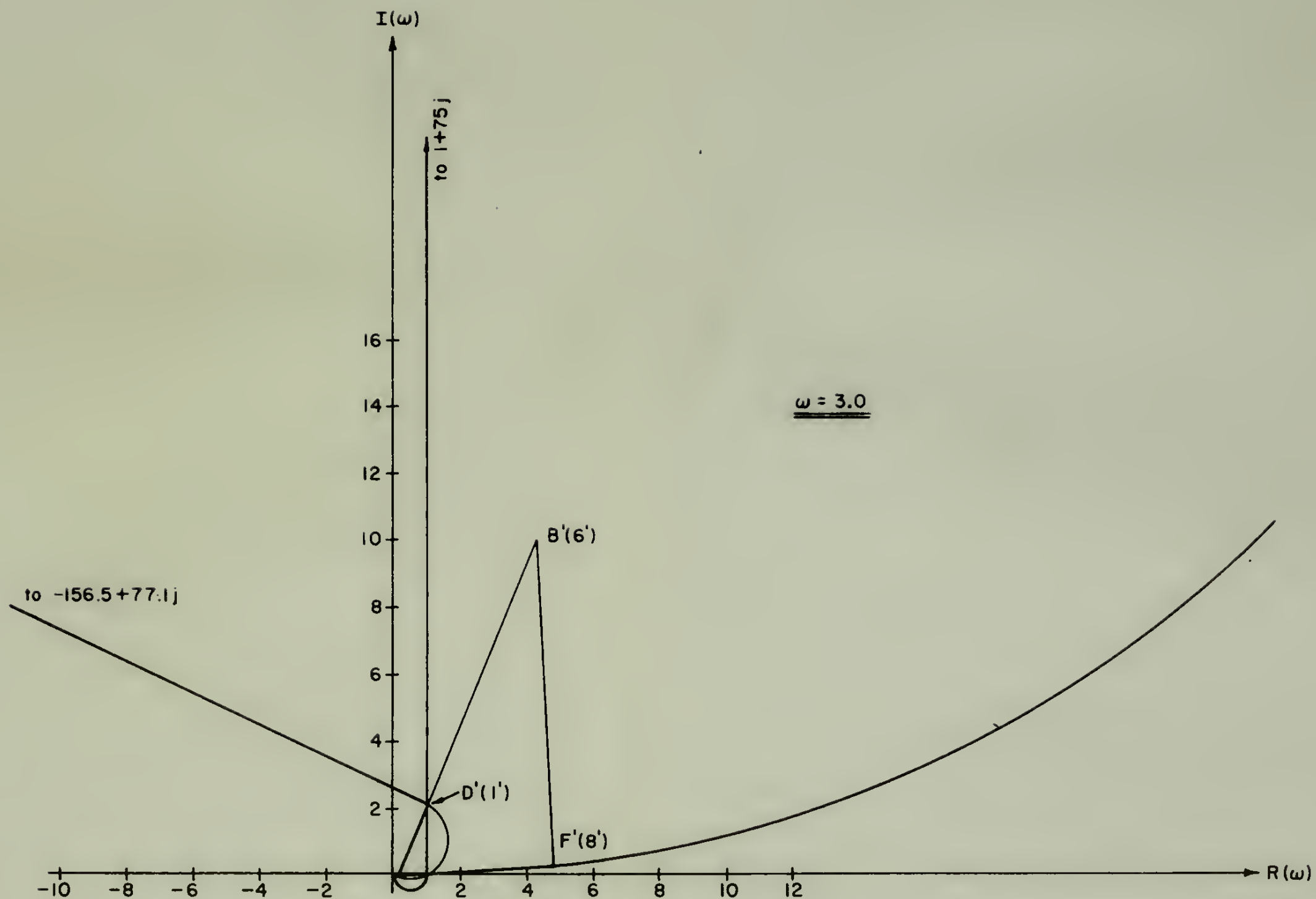


Fig. 4-22: Variable Parameter Mapping

Thesis
S557

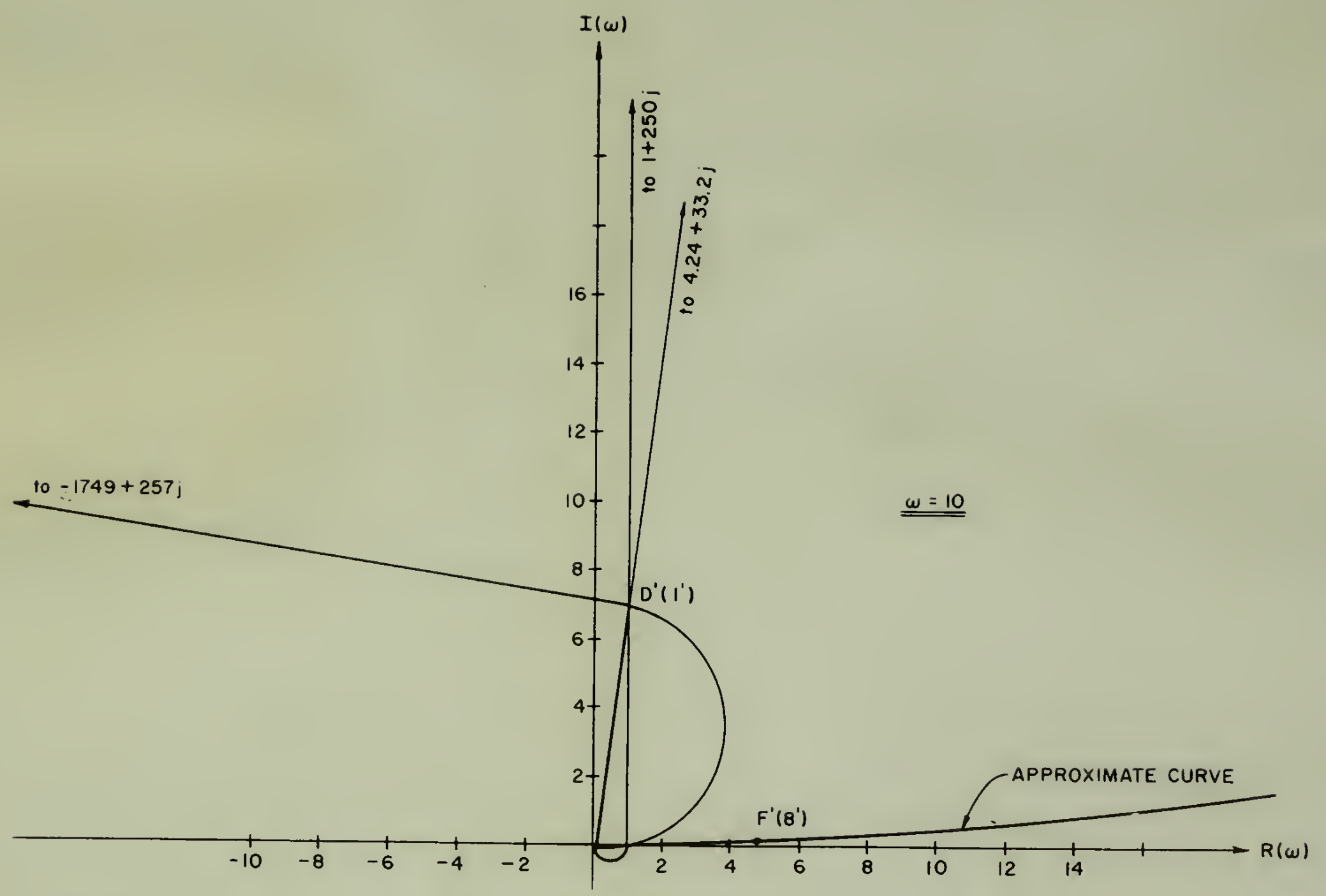


Fig. 4-23 : Variable Parameter Mapping

thesS557

Stability and performance of manned cont



3 2768 001 00610 9

DUDLEY KNOX LIBRARY